

1 ATGAGCCAGC CCAGGCCCCG CTACGTGGTA GACAGAGCCG CATACTCCCT
51 TACCCCTCTC GACGATGAGT TTGAGAAAGAA GGACCGGACA TACCCAGTGG
101 GAGAGAAACT TCGCAATGCC TTCAGATGTT CCTCAGCCAA GATCAAAGCT
151 GTGGTGTGG GGCTGCTGCC TGTGCTCTC TGCTCCCCA AGTACAAGAT
201 TAAAGACTAC ATCATTCCCTG ACCTGCTCGG TGGAACCTAGC GGGGGATCCA
251 TCCAGGTCCC ACAAGGCATG GCATTTGCTC TGCTGGCCAA CCTTCCTGCA
301 GTCAATGGCC TCTACTCCCT CTTCTTCCCC CCTCCTGACCT ACTTCTTCTC
351 GGGGGGTGTT CACCAAGATGG TGCCAGGTAC CTTTGCCGTT ATCAGCATCC
401 TGGTGGGAA CATCTGCTG CAGCTGGCC CAGAGTCGAA ATTCCAGGTC
451 TTCAACAATG CCACCAATGA GAGCTATGTC GACACAGCAG CCATGGAGGC
501 TGAGAGGCTG CACGTGTCAG CTACGCTAGC CTGCTTCACC GCCATCATCC
551 AGATGGGTCT GGGCTTCATG CAGTTGGCT TTGTGGCCAT CTACCTCTCC
601 GAGTCCTTCA TCCGGGGCTT CATGACGGCC GCCGGCCTGC AGATCCTGAT
651 TTCGGTGCTC AAGTACATCT TCGGACTGAC CATCCCCCTCC TACACAGGCC
701 CAGGGTCCAT CGTCTTACCC TTCAATTGACA TTGCAAAAAA CCTCCCCCAC
751 ACCAACATCG CTCGCTCAT CTTCGCTCTC ATCAGCGGTG CCTTCCTGGT
801 GCTGGTGAAG GAGCTCAATG CTCGCTACAT GCACAAGATT CGCTTCCCCA
851 TCCCTACAGA GATGATTGTC GTGGTGGTGG CAAACAGCTAT CTCCGGGGC
901 TGTAAGATGC CCAAAAGTA TCACATGCG ATCGTGGGAG AAATCCAACG
951 CGGGTTCCCC ACCCCGGTGT CGCCTGTTG CTCACAGTGG AAGGACATGA
1001 TAGGACACAGC CTTCTCCCTA GCCATCGTGA GCTACGTAT CAACCTGGCT
1051 ATGGGCCGA CCCTGGCCAA CAAGCACGGC TACGACGTGG ATTGAAACCA
1101 GGAGATGATC GCTCTCGGCT GCAGCAACTT CTTTGGCTCC TTCTTTAAAA
1151 TTCATGTCAT TTGCTGTGCG CTTCTGTCA CTCTGGCTGT GGATGGAGCT
1201 GGAGGAAAAT CCCAGGTGGC CAGCCTGTGT GTGTCTCTGG TGGTGATGAT
1251 CACCATGCTG GTCTGGGGG TCTATCTGTA TCCTCTCCCT AAGTCTGTGC
1301 TAGGAGCCCT GATCGCTGTC AATCTCAAGA ACTCCCTCAA GCAACTCACC
1351 GACCCCTACT ACCTGTGGAG GAAGAGCAAG CTGGACTGTT GCATCTGGT
1401 AGTGAGCTTC CTCTCCCTC TCTTCCTCAG CCTGCCCTAT GGTGTGGCAG
1451 TGGGTGTCGC CTTCTCCGTC CTGGTCGTGG TCTTCCAGAC TCAGTTTGA
1501 AATGGCTATG CACTGGCCCA GGTCACTGGAC ACTGACATT ATGTGAATCC
1551 CAAGACCTAT AATAGGGCCC AGGATATCCA GGGGATTAAA ATCATCACGT
1601 ACTGCTCCCC TCTCTACTTT GCCAACCTAG AGATCTTCAG GAAAAGGTC
1651 ATCGCCAAGA CTGTCTCCCT GCAGGAGCTG CAGCAGGACT TTGAGAATGC
1701 GCCCCCCCAC GACCCCCAAC ACAAACAGAC CCCGGCTAAC GGCACCAAGCG
1751 TGTCTTATAT CACCTTCAGG CCTGACAGCT CCTCACCTGC CCAGAGTGAG
1801 CCACCAAGCCT CCGCTGAGGC CCCGGGGAG CCCAGTGACA TGCTGGCCAG
1851 CGTCCCACCC TTGCTCACCT TCCACACCCCT CATCCTGGAC ATGAGTGGAG
1901 TCAGCTTCGT GGACTTGATG GGCATCAAGG CCTGGCCAA GCTGAGCTCC
1951 ACCTATGGGA AGATCGGCCT GAAGGTCTTC TTGGTGAACA TCCATGCCA
2001 GGTGTACAAT GACATTAGCC ATGGAGGCCT CTTTGAGGAT GGGAGTCTAG
2051 AATGCAAGCA CGTCTTCCC AGCATACTG ACCGAGTCT CTTTGGCCAG
2101 GCAAAATGCTA GAGACGTGAC CCCAGGACAC AACTTCAAG GGGCTCCAGG
2151 GGATGCTGAG CTCTCCCTGT ACGACTCAGA GGAGGACATT CGCAGCTACT
2201 GGGACTTAGA GCAGGGAGATG TTGGGGAGCA TGTTTCACGC AGAGACCTG
2251 ACCGCCCTGT GA (SEQ ID NO:1)

FEATURES:

Start Codon: 1
Stop Codon: 2260

FIGURE 1A

HOMOLOGOUS PROTEINS:

Top BLAST Hits:

		Score	E
gb AAF81911.1 AF279265_1	(AF279265) putative anion transpor...	476	e-133
gb AAF71715.1 AF230376_1	(AF230376) prestin [Meriones ungu...	471	e-131
ref NP_000432.1	pendrin [Homo sapiens] >gi 11421915 ref XP...	451	e-125
ref NP_035997.1	Pendred syndrome homolog (human); Pendred'...	448	e-124
ref NP_062087.1	Pendred syndrome homolog (human) [Rattus n...	447	e-124
ref NP_067328.1	down-regulated in adenoma [Mus musculus] >...	434	e-120
ref NP_000102.1	down-regulated in adenoma protein [Homo sa...	418	e-115
sp 070531 DTD_RAT	SULFATE TRANSPORTER (DIASTROPHIC DYSPLASI...	365	1e-99
ref NP_000103.1	sulfate anion transporter 1; Diastrophic d...	362	1e-98
ref NP_031911.1	diastrophic dysplasia [Mus musculus] >gi 2...	357	4e-97

BLAST to dbEST:

	Score	E
gi 8630793 /dataset=dbest /taxon=960...	523	e-146

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

Expression information from BLAST dbEST hits:

gi|8630793 Human head-neck

Expression information from PCR-based tissue screening panels:

Human fetal lung

FIGURE 1B

1 MSQPRPRYVV DRAAYSLLF DDEFEKKDRT YPVGEKLRNA FRCSSAKIKA
51 VVFGLLPVLS WLPKYKIKDY IIPDLLGGLS GGSIQVPQGM AFALLANLPA
101 VNGLYSSFFP LLTYFFLGGV HQMVPGTFAV ISILVGNICL QLAPESKFQV
151 FNNATNESYV DTAAMEAERL HVSATLACLT AIIQMGLGFM QFGFVAIYLS
201 ESFIRGFMTA AGLQILISVL KYIFGLTIPS YTGPGSIVFT FIDICKNLPH
251 TNIAISLIFAL ISGAFLVLVK ELNARYMHKI RFPPIPTEMIV VVATAISGG
301 CKMPKKYHMQ IVGEIQRGFP TPVSPVVSQW KDMIGTAFSL AIVSYVINLA
351 MGRTLANKHG YDVDSNQEMI ALGCSNFFGS FFKIHVICCA LSVTLAVDGA
401 GGKSQVASCL VSLVVMITML VLGIYLYPLP KSVGALIAV NLKNSLKQLT
451 DPYYLWRKSK LDCCIWVVSF LSSFFLSLPY GVAVGVAFSV LVVVFQTQFR
501 NGYALAAQVMD TDIVVNPKTY NRAQDIQGK IITYCSPLYF ANSEIFRQKV
551 IAKTVSLQEL QQDFENAPPT DPNNNQTPAN GTSVSYITFS PDSSSPAQSE
601 PPASAEAPGE PSDMLASVPP FVTFHTLILD MSGVSFVDSL GIKALAKLSS
651 TYGKIGVKVF LVNIHAQVYN DISHGGVFED GSLECKHVFP SIHDAVLFAQ
701 ANARDVTPGH NFQGAPGDAE LSLYDSEEDI RSYWDLEQEM FGSMFHAETL
751 TAL (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 3

1	153-156	NATN
2	156-159	NESY
3	580-583	NGTS

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 2

1	45-47	SAK
2	445-447	SLK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 11

1	18-21	TLFD
2	158-161	SYVD
3	240-243	TFID
4	365-368	SNQE
5	459-462	SKLD
6	556-559	SLQE
7	635-638	SFVD
8	691-694	SIHD
9	722-725	SLYD
10	726-729	SEED
11	732-735	SYWD

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

Number of matches: 2

1	7-15	RYVVDRAAY
2	447-454	KQLTDPYY

FIGURE 2A

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 10

1	77-82	GGLSGG
2	78-83	GLSGGS
3	89-94	GMAFAL
4	103-108	GLYSSF
5	335-340	GTAFL
6	435-440	GALIAV
7	481-486	GVAVGV
8	485-490	GVAFSV
9	581-586	GTSVSY
10	681-686	GSLECK

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	51	71	0.893	Putative
2	82	102	1.020	Certain
3	107	127	1.729	Certain
4	130	150	1.497	Certain
5	186	206	1.723	Certain
6	228	248	1.517	Certain
7	256	276	1.898	Certain
8	288	308	1.252	Certain
9	338	358	1.568	Certain
10	383	403	1.304	Certain
11	412	432	2.345	Certain
12	469	489	1.997	Certain
13	619	639	1.146	Certain

FIGURE 2B

BLAST Alignment to Top Hit:

>gb|AAF81911.1|AF279265_1 (AF279265) putative anion transporter 1 [Homo sapiens]

Length = 738

Score = 476 bits (1224), Expect = e-133
 Identities = 263/724 (36%), Positives = 428/724 (58%), Gaps = 36/724 (4%)
 Frame = +3

Query: 54 LFDDEFEKDR--TYPVGEKLRNAFRCSSAKIKAVVFGLPVLSWLPKYKIKDYIIPDLL 227
 Sbjct: 15 L + EE R + P + R +CS A+ A++ LPVL WLP+Y ++D++ DLL 74

Query: 228 GGLSGGSIQVPPQGMAFALLANLPAVNGLYSSFFPLLTYYFLGGVHQMVPGTFAVISILVG 407
 Sbjct: 75 GLS +Q+PQG+A+ALLA LP V GLYSSF+P+ YF G + GTFAV+S++VG 134

Query: 408 NICLQLAPESKFQVFNNATNESYVDTAAMEAERLHVSATLACLTAAIQMGLGFMQFGFVA 587
 Sbjct: 135 ++ LAP+ A N+S ++ A +A R+ V++TL+ L + Q+GLG + FGIV 186

Query: 588 IYLSESFIRGFMTAAGLQILISVLKYIFGLTIPSYTGPSSIVFTFIDICKNLPHNTIASL 767
 Sbjct: 187 YLSE +RG+ TAA +Q+ +S LKY+FGL + S++GP S++T ++C LP + + ++ TYLSEPLVRGYTTAAAVQVFSQLKYVFLHSSHSGPLSLIYTLEVCKLPQSKVGT 246

Query: 768 IFALISGAFLVVKELNARYMHKIRFPIPTEMIVVVVATAISGGCKMPKKYHMQIVGEIQ 947
 Sbjct: 247 + A ++G LV+VK LN + ++ PIP E++ ++ AT IS G + ++ + +VG I VTAAVAGVVVLLVKKLNDKLQQQLPMPPIPGEELLIGATGISYGMGLKHRFEVDVGNIP 306

Query: 948 RGFPTPVSPVVLQWKDMIGTAFSLAIVSYVINLAMGRTLANKHGYDVDSNQEMIALGCSN 1127
 Sbjct: 307 G PV+P + ++G+AF+A+V + I +++G+ A +HGY VDSNQE++ALG SN AGLVPPVAPNTQLFSKLVGSFTIAVVGFAIAISLGKIFALRHGYRVDNSQELVALGLSN 366

Query: 1128 FFGSFFKIHVICCALSVTLAVDGAGGKSQVASLCVSLVMITMLVLGIYLYPLPKSVLGA 1307
 Sbjct: 367 G F+ + C+S +L + GG SQA SL +++ ++ LG + LPK+VL A LIGGIFQCFPVSCSMSRSLSVQESTGGNSQVAGAISSLFILLIIVKLGEFLHDLPKAVLAA 426

Query: 1308 LIAVNLSKNSLQLTDPPYYLWRKSKLDCCIWVVSFLSSFFLSPYGVAVGVAFSVLVVVFQ 1487
 Sbjct: 427 +I VNLK L+QL+D LW+ ++ D IW+V+F ++ L+L G+ V V FS+L+VV + IIIVNLKGMLRQLSDMRSLSWKRANRADLLIWLVTFTATILLNLDLGLVVAVIFSLLLVVVR 486

Query: 1488 TQFRNGYALAQVMDDTDIYVNPKTYNRAQDIQGKIKIITYCSPLYFANSEIF----- 1637
 Sbjct: 487 TQ + L QV DTDIY + Y+ A++++G+K+ + +YFAN+E + TQMPHYSVLGQVPDTDIYRDVAEYSEAKEVRGVKVFRSSATVYFANADEFYSDALKQRCGV 546

Query: 1638 -----RQKVIAK--TVSLQELQQDFE-NAPPTDPNNNQTPAN-GTSVSYI----- 1760
 Sbjct: 547 ++K++ K + L++LQ++ + P N TS+ + DVDFLISQKKLLKKQEQQLKLKQLQKEEKLRKQAASPKGASVSINVNTSLEDMRSNNVED 606

Query: 1761 -----TFSPDSSSPAQSEPPASAEAPGEPSDMLASVPPFVTFTLILMSGVSFVDLMGI 1925
 Sbjct: 607 S D A + ++AP + S + A P FH+LILD+ +SFVD + + CKMMQVSSGDKMEDIATANGQEDSKAP-DGSTLKALGLPQPDFHSLILDLGALSFVDTVCL 665

Query: 1926 KALAKLSSSTYKGIGVKVFLVNIHAQVYNDISHGGVFEDGSLECKHVPSIHD AVLFAQAN 2105
 Sbjct: 666 K+L + + +I V+V++ H+ V + + G F D S+ KH+F S+HDAV FA + KSLKNIFHDFREIEVEVYMAACHSPVVSQLEAGHFF-DASITKKHILFASVHDATFALQH 724

Query: 2106 ARDV 2117
 R V
 Sbjct: 725 PRPV 728 (SEQ ID NO :4)

Hmmер search results (Pfam):

Model	Description	Score	E-value	N
PF00916	Sulfate transporter family	254.5	1.5e-72	1
PF00189	Ribosomal protein S3, C-terminal domain.	3.3	8	1

FIGURE 2C

Docket No.: CL000861CON
Serial No.: TO BE ASSIGNED
Inventors: Karl GUEGLER et al.
Title: ISOLATED HUMAN TRANSPORTER...

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00916	1/1	187	497	..	1	328	1.5e-72
PF00189	1/1	651	661	..	79	89	3.3

FIGURE 2D

1 CTGGGTTCCCT ATGTGGGGAG GTCATGCTCC CCACTCATTG AGCCCCCCC
51 GGCACAAACAC CTGGACAGCC AGACCCATGC AGACTCTGGA GCAGGTGGAG
101 AGGAAGAGTG AGACCACCCC GCCTCACGGG CGGTGAAGGG CCGGCAGCCT
151 CTGAATAGTC TCTGCTAGGA GGTAGAAAAGC ACCCTCCCAT CTTAATCATA
201 GTAATCATCG CCACTACCAT TTACTGGGTG CCTATAAAAG GCCAGCCTCT
251 TCATAACACAT GATCTCACTG AATCCTCATA GCATCTGCCT GCGACTGTTA
301 TTATCCCCAT TTACAGATGA AGAAACTGAA TCTTGAACC CAGGTCACT
351 GGCTCTCAAA CTTGTGCTGT TTTCCCTAAG CCACCCGGTC TCTCATTCT
401 CCCACTGAAA TGTCTCACAT GCCATTGCCC TTACTCATTT CTGCCCAGT
451 CTCCCTCCAAA ACACCATTTA TCAATTGCGT CAACAAGTAT GTGTTGAGTA
501 CACACTAAGG GCCAGGCAGG GGGCTGGCA CAGGCCTGTTG GGGTAGGTT
551 ATTCTCCAC CTTCGCTTCT GCTGGGTATC ACCTGTGGGG TCTTGCCGGG
601 CATCCCACCC TCACCTGTAG TTCAAGTGGA CCTTGGGATC CCAAGACCAA
651 ATGAATGGAA TGCACCCAGCC CAGCCTTCAC CAACCTTGAGC ACAATCTTAT
701 TCATAATAGA AACTCACATT TGATCACAC TTTACATTTT ACACAACCCC
751 TTCTTATCCA TTAACTCATT TGATCTTCAC AACAACCTG TGAGATATGT
801 CTGTTACTCC CACTTTAGTG ATACAGAAC TGAGGTTTGA AAAGTAATGC
851 TGACCATTTCT GCCTCATTAA TAAAAGCAGG ATTAAACCCAG GCTCCCTGGAC
901 CCTTCCACAA AAGGCATTAA GCAACCTGCT CCCCTCTGAC AACCTCCCT
951 GTCACCCAGG CTCTCCCTG GGAAGTTGGG GGCATCTCTA GCCCCCAAGT
1001 AGTTACTCAT TTTCAACCCCC ATCTCAAATC TTTTGCCTAA CTGGCCACAG
1051 CCACCCACCA CTCCCCACCT CCCAGATACA AATCCTCACT CTAAGCCTTC
1101 CCCATCTCTT TCTTCTCTGT CCTTCTTCTT CTGTGGTCCT CTGAGCAACT
1151 TCTCCCAAGCT CTGGGAGGTG GAGGGGAGGT GGGAGACCCA GTAAATTGGAA
1201 GAGGGAGGGG GAAAGGTTCC TACAGGGAAC TCTCTGGGGC CTCAGGGGCC
1251 CTGGCACTCA GCTCTGCCCA TCTCAGCTCC TGGAACGTCA GCCAGGTTGC
1301 GCAAAAAGTG AGGAGGAGAG GAGCAGCAGT ACACAAGGGT GGGGGAAAGA
1351 TTAGGCACAG GAAGCCGTGG GAGAGAGAGC CGGCAGGTGG ACCATCCTGG
1401 TTTCCCCACA CACACCATTG TCCCCCTGGG AAACCTGTTG GTGAAGTTCT
1451 AGATGTCTT TCCAAGAAGG GTCTCTTCA GGTCTATCTCA GCTATCCCC
1501 TGCTCTTAGG CAAGCTGTTT TCTGTTTCTT CCAAGCTGAC TGGCTGAATG
1551 GTAGGAGCCT TTCTGCCAGG GAAACTAAGG TCTGGGAAGG GAGTATGGCT
1601 TGTGGGGACA CCAGGGGTCA GGGGAGGGGA GGGTCCACCT GCTGAATCAA
1651 GTGGGGCTC CTGCCCTCGT GATTCCCCCTT TGCCCTGGTGC TCAGTGGGG
1701 TGATGGTGAC GCCACAGGTG TGGAGTGCCA GCCACGTGCT GAGCGCCAAG
1751 CAAAACAGCC AGGGTGTAGTC TATGCATCAT CAGTGCCTGG GAAGGAAGGC
1801 CACTGGCAGC AGGGAGTCTG ACGGAAAAAAC TTGACAGAGG GAAGGGAGGC
1851 ACCTTGCTT ATCGGGCGGG GGAAGGCCAG AATAAAACTC TGCTACTGCA
1901 AGGACCAAGAG AGAGAAGGCC TGGGCTGGCA CTAGGGAGGG ATGTTCCCTC
1951 ACCCTCCCT CCTCTGCTTC TCCCAAAGCT TGAAATGCC CCAGATATGA
2001 GCCAGCCCAG GCCCCGTAC GTGGTAGACA GAGCCGCATA CTCCCTTACC
2051 CTCTTGACG ATGAGTTGA GAAGAAGGAC CGACATACC CAGTGGGAGA
2101 GAAACTTCGC ATAGCCTTCAG GGTAACTGGT CCAGAGGCCA GACTTCTGCC
2151 TCCTCTGCTC CCTACCAAAA TCCCTTCTGC ACCAGGACAC GGCTTCTGCA
2201 CTGGTATCCC TAAGATGGGG TTAAGGGAAG CCTCTGGGGAA GTGAGGTTCT
2251 GAATGATGAA TTAAAGATCC TACAACCTCA TCTGTACTGA GACCCCCCAGG
2301 GAGGATGGGG AGCAGGAGCA AGAACCATCC AGAAGGGTTA TATGGCATT
2351 CCAAACCCCT GCATGGCATC TCCCATATTG TCAATTCAAC CGGGTCTCTC
2401 TGGGTTTGTGTT AAGGCATGGT AGATGAGCAT CTACGTTATG GAGGGGTGGG
2451 GAGCATCAGA GCCCTTACTCT CATGCCCTGT TCCCTCCCTTA CAAAAAAATAC
2501 CTGAAGTTAC CATCACCCCCA GGTCTTTGT CCTTCTCCCTC CGGGATGTT
2551 CTTCTCCAC TTGGTCCAGA GAATGCCAA AGGAGGCCCT AAATTCTGAA
2601 ACTTTCTGAA GGGGACCTAC CAGGGTGTAG TCTTACCAAGC GCCCAGGGTC
2651 TTTCCACTCT CATCTCCCTG GAAATGCGAT GGTGGGTATG AAACCTTGT
2701 CCTAAGTAGG CGCTACACAA GGTGATCCAT ACCCACACCC CAGGAGGCTG
2751 GGGCTGCGGG TGTCACCCCTC CCCATTCCCA GACTCCTGGC AGACCTCCTC
2801 TGGCCCAAGCT ATAGGCCAAC TCACTCTCCC TCACTCCCTT GGGGAAACGG
2851 CTGATTCACT TACCTGGATT GAGGTCACTG GCAATGGCTG AAGTGGAGAC
2901 GCAGGGTGGAA CTGGTTCAAGG CCGGGGGAAAT CACCCACTTG AGTTTGACT
2951 AAAAGCCCCA GCCCAGCCCT GTTTCTCTTGG GAGGCTCCA TTTCTGCCA
3001 GTTACAGTCT GTCTCACAG CTGTGCTCT CAGACAGGTG GTCTCTGCCA
3051 GTCTTTGTGC CCAAGACTTT AGGGCACAAA GTCTGAGGAT GAGAAGATCT
3101 GCTATTGTC TAAAAGATTA GGATAATGAA AGCTGTAAAG GGATATAGCA
3151 AACTAACAAAT TCCTATGATA CTGGCATGAG AGCCTTGAAC AGTGCCTGGC
3201 ATAGAGAAGG TGCACCAATA AATATTGTT TCATGAATGA ATGAATGAAT
3251 GAATGTCTAG AAAGCTAATC CCTCTCAGCC TCTGTTTCA GTTCTCTT
3301 CAAGCTTCAG ATTGCTTTGC CCAACATACA GCAGACTTGC AAGTAAGGTT
3351 GGGCATGGAC TAGCCCTCAA ATGAGTTGTT TTTCTTCCC TAGCCAGCTC

FIGURE 3A

3401 TCTATTCTATA AGTCCGGCTT TCTCTGCCAC AAACAGACCT GATGGAGCCC
3451 CTGCAAGGGCT GGTTCCTCTCT TCAAGCAAGG CTTTAGAGTT GCATTAAGCA
3501 ATTTATCCCC CGTCCACCTC CCCTTCAGC ATCCCAGGGA TGGCAGAGGC
3551 ACCCATGAGC CCCAGAAGGG ACAGGGGTA AGATATTGAT GATGATGCTT
3601 TTTCTTGGAG TGTTAGTTGG AAGAGAAAAT CTGCCCAGAC TTTCCAAGGT
3651 ACAAAAGCATT GTCTTTGTTG GTTTCAGTCT TGTTGACAT CCAGGGGACCC
3701 GAGTGTCAAGG GAAACTATTG TTGAGCAAGA GCAAAGAGCA GGAATTGGTG
3751 CTGGCAGGA AAGGAAGCCT CATCAGAGCA GGGCAGTGAG TCACCAAATG
3801 GCCCTTAAGT ATTTGAGTTC CCTCAACTGG GAGAAGGAAA GCAAATGCC
3851 CTCACCCACT TCCAGTCATC AATCCACCGG CTGTCACCCCT TGAGTTTGT
3901 AGCCCTGTT CCTACCGCTC CTGAGTTCT ATGAAAGGAC CTTGAGGTGT
3951 TCAACAAACA GGGAAAGGGAT CAACTCTCCC CACCCCTGCGT TGACCAATGA
4001 ATTCTTCCCT CCTCTGCTGC CCAGTGAATT AACAGGAGAA AGAACTCCGG
4051 TATGGAGTT ACCACACATA AAGGATAGTG AGTCAGCAGA GTGCACCTG
4101 CAGGAAACAT AGAGCCTTCC TTTTCAAGGA ATTTCTAAGA AAAATGGCAG
4151 CAGGCAGGCC CCACTCGGGT GTATTCACTC ATTCACTTTAT TCAACAAATA
4201 TTTACTAAGT GCCCCCTGTC AAGGCTCGAG GTGTACAAAG ATGAACAGGA
4251 GAGCTAGACT TCTTGCCATG CGTGGTGGGG TTTGCTGCCT AGTGGGAGAG
4301 ACAGACAAAA AGCAAGGAAT GCACACACAG GATGCACACA CAGCGGCAGG
4351 AACCAAGGTG CAGTTACCCA GGCCTGGAT CAGACAGACAGA GGACTCAGAG
4401 GAGACTTCC CAGAGAAAAG CCATCTGAGC CAAGGGATGG ATCTGATACC
4451 TCCGAAGGCT GAGCCACCAT AACACTCATC CCTTTAAGCC AAGTCTTATA
4501 AACTCCCCAG DTAAAGCAGG GGCAGTCAGA AGACCTCCAG CTAATGCCCA
4551 GGACAAGTTG ATGAGCTCTC AAGAAAAAGT TCCTGCCTTT TCTTCTCAAT
4601 ATCCCTGGCA CACAGTTCA TGAACTTTGA ATGAACCAAT GAATGAAATG
4651 AGCAGGATAT GATAATCCCT CTCCAACACG GAATGTCCAA GCCATGCAGA
4701 GCCGACTTGG AATTTTCCCC GTTCCCTTCC AGATGTTCCCT CAGCCAAGAT
4751 CAAAGCTGTG GTGTTGGGC TGCTGCCTGT GCTCTCCCTGG CTCCCCAAGT
4801 ACAAGATTAA AGACTACATC ATTCTTGACCC TGCTCGGTGG ACTCAGCGGG
4851 GGATCCATTC AGGTCCCAACA AGGTGAAGGG GCTCCTTCAG CCAGGCCTGG
4901 ATTGCCACTC CCCTCACCAT TCCTCTCCTC ATCCCCACTC CATCCCTCTG
4951 TGATCCCCAT AAGCTAGTCA TGCTGCTGAG CTTCAGTCTC GTTGCCTCT
5001 GCAGGCATGG CATTGCTCT GCTGGCAAC CTTCCTGCAG TCAATGCC
5051 CTACTCCCTCC TTCTTCCCCC TCCTGACCTA CTTCCTCCTG GGGGGTGTTC
5101 ACCAGATGGT GCCAGGTAAG GCCTCTCCC TCTGGGCAGG CAGGATGACC
5151 CAGACACAAA GGATGGGAGG TGTTGAAAG GGGCCTCGGG AGATTTCCA
5201 TCTGCATTCT CTTGGAGTTG TCCCTGGTCA GTCTCTAGGGG AATGGTCACT
5251 GTGAATGTCA TTTCCAGGTC CTCGGTGACC TTGGAGAAAC CACTGAGCCT
5301 CTTTGAGTTG AGTTAGCATT ACCTGTTCCA TCTTCCTCCT AGGAATGAGA
5351 GGAAGACTTA GCAGAACAAAG ATATACCATA TGCTATAACA TGCTTAAACA
5401 GATGTGAGAA ATCACCATC AACTCCCTGG TTGGTCCCAAG CGGCCACTA
5451 CAGGGACATT TGGACTTCTC TGTTGCTAAG TGAGATGGAG GAAAGCCTGG
5501 TCACAGGGC TTGTTCTGG TTCAAGCTCT GCTTATATT CTTATTCTG
5551 AGTCATTTT CTCACGTGTC CTGTATGACA ATATTGACCA TTGGGGTAAA
5601 AGCACCTTGA AAAGCATAGA TCATGGTTAG AGTGAGTGGT TGTTATTATT
5651 GTGTTGGAGA AGAGCCTTGG AGGTGCAGGG ATCCATCCCC CTGGGGTCTG
5701 GAAGCATTCC TGGGCCCTT TCTGGTTCC ATCGGTGTGG TTCAAACCTC
5751 TGATTTTGC TGGCTGGGT GGGCACACA GGACCTTTG CCGTTATCAG
5801 CATCTGGTG GGTAACATCT GTCTGAGCT GGGCCCAGAG TCGAAATTCC
5851 AGGTCTCAA CAATGCCACCC AATGAGAGCT ATGTGGACAC AGCAGCCATG
5901 GAGGCTGAGA GGCTGCACGT GTCAGCTACG CTAGCCTGCC TCACTGCCAT
5951 CATCCAGGTG AGGGGGCAGC CCCCAACCC GCTAGAAGGG CATCAGACCA
6001 CCCTGCCCTC CCCTCAAAGC CTTAGTTTG ATGCTAAATC TGATTTAGGG
6051 GGCTGGGTGT GGAGGCTCAT GCCTGTAACT CCAGCACTTT GGGAGGCTGA
6101 GGAGGGTGGG TCACTTGAGG TCAGGAGTTT GAGACCACT TGACCAACGT
6151 GATGAAACCC CATCTCTACC AAAATAACAA AAAATAATCCA GGCTTGGTAG
6201 TATGCGCTG TAGTCCCACCC TACTCAGGAG GCTGAGGCAG GAGAATCACT
6251 TGAATCCGGG AGGCAGAGGT TGCACTGAGC TGAGATCGCG CCACTGCACT
6301 CCAGCCTGGG TGACAGAGGC AGACTCCGTC TCAAAAAAAA AAAAAAAA
6351 AAAAAAAA CCAAGTTAG GGCTCACCTC CTCCCTCCTC CCCATCCCC
6401 GGCTAAAGTG AACCTTGAAA ATTAACAGTA TCTCTCTCATC TGATGTAGC
6451 AGGACCATAC AAAAAAAACAA CAGCTGTACC TGTTAAACT GTCCTGAGCT
6501 TAAACCTGT AAAAGACTCA CAGCCTCTCT CCATTATCCC GTGGAGAAAC
6551 CCAACTCTCT GCCAGCATAG TCTTGAGAC TGCTAATTTT CTCTAACATC
6601 CCTCACTCCG CTCCAGCCTC CTCTGCTCCA AGCCACAGCA GCAGTTGCAC
6651 AACATAAAATT GAGCTTCTGC AAATGGTTGC AAAGGATTCT GCTAGTTTT
6701 ATGAAGGGAA GCACAAACATG ACAGAAATGCA AGAGCAAAAC ACAGTCCAG
6751 AGAGGCCCTT TTCAATTCTC CATTCTATTG TGTTTGTGCC AAGAACTAGG

FIGURE 3B

6801 CTAACCCCTG GGATACAAAG ATAAGTAAGA AAGAGGTCCA ATTACAAGT
6851 TGCTCACAGC CCAGCAGG AAGGAGCCAT GTCAACAGAT AAATTTGTAT
6901 GCAGTGAGAT AAGCAGAA GTAGAGCCAT GTCAAAGAC TGTAGGGACA
6951 CAGAGCAGAG TCACGGAGGA CCTCAAAGAG GAGGTGACAC TCCACCTCTC
7001 TAAAGGATG AGAACTTAAC CAGGAACAAG GTATAACAGAG GATGGTCAG
7051 GCAGAAGGGA ACAGTGCCTA AAAACACTGA GGCCTGAGAG AGTGTGATCT
7101 GCGCAGGCAA AGTAAGGGGC TTGGTGTGGC TGAGGGTAG AGGGCCCAGA
7151 AGAGGATGGA AAAGTAGGCA GGAGCCAGAC AATGAGATCT GGGGTCTGTT
7201 CTCTGACAGC GACTTTGGG CTGATTGGC GTTTATAAGG ATCGTTTGGG
7251 CTACAACTG ATGAGTGGG GGTGGATTAG AATCAAGGCA GGGGACCTGT
7301 TGGGAGACTC TGCAAGGGCC CAGGCAGGAA TAATGCAGGC GAAGACCAGG
7351 TAGAGAAAGA GATGGGGCTG GACTGAAAAA GAATGTTTA CCAGGAGCTT
7401 GGTGATAGAC TGGATGTGGG AGGTAAAGGA GGATGACTCT CAAGTTTTG
7451 GTTGGGCAAC CAGGTTAATG ATGGTGTAT TTACTGAGAG AGAAAACACT
7501 GGGGAGGAC TAGACTTATT TTACAGATAA GCAAAGCCA GAGAGGTGAT
7551 GTGACAGAAA GGCCCATGCT CTAAAGGAGC TGAAGGTCTG ATGGCAGCCA
7601 TGTAGAGCAC AGTGAAGGGC AGGTGAAGGT CACAGATGGT CCAATTCCCT
7651 CAAGCTACTG CTACGCTAGG ACTGCACGGA GCTCCAGACC TCGTGTGTT
7701 TGGGGCGGGT CGTTGGAACT GCTGAACAC ATTGGTCTTC CGCCACCAAC
7751 CACCTTTTC CTCCCTCTAG ATGGGCTGG GCTTCATGCA GTTTGGCTTT
7801 GTGGCCATCT ACCTCTCCGA GTCCCTCATC CGGGGCTTC TGACGGCCGC
7851 CGGCCTGCAG ATCCTGATT CCGTGTCTAA GTACATCTTC GGACTGACCA
7901 TCCCCCTCTA CACAGGGCCA GGTCATCG TCTTGTGAG TCTGGGGATG
7951 CACCCCTGCC ATTGGAGCAA GGCTCCAGCA GACACATGAG GAGGATGTAC
8001 TGTTTAAGA TGTGTGAGC TCCTCATTCG AAGGGCTGGC TTAGCTGTTG
8051 TTCAGAGAGG ATTCTGAGGG GGTTCTGTC TTGGGAGGGT CAAAGTCATG
8101 ACTCACAGAG GTTCTTGGT GTAAATACCT GCAGAAAAGA GCTGTACATT
8151 CTCCGCCAGT TCCCCATTCT AGTGCCTCAA CCCTCCCTG CCTGGAAGT
8201 CCTGCCCTAT GTCTAATCTC CATCCCCCTC CCTTCAGGCC AAACCTTTCT
8251 AAAGAAAAAG AAAGCATTCC TTTTCTAGCA CAAGTCCCC ATGTGCCCTT
8301 TGGGAAAGGG CGGTGGGCGA CGGGACAGGG TTCTGATCA GGGTTTTAAT
8351 TCTGTCTGG TGTGCTCTA TTAGCTTGA TGGCATCCCT TCCCTGGTC
8401 AGACACCCAA AGGTGGGGTA TTATGGAAG AAGGGGTGGG AGCCTGTGAG
8451 CATGATGCTC TTTCCCCCAG ACCTTCATTG ACATTTGCAA AAACCTCCCC
8501 CACACCAACA TCGCTCGCT CATCTCGCT CTACATCAGCG GTGCCCTTCT
8551 GGTGCTGGT AAGGAGCTCA ATGCTCGCTA CATGCACAAG ATTGCTTCC
8601 CCATCCCTAC AGAGATGATT GTGTAAGGA CCTTGTTCAG AGCTGGGATG
8651 TTGGGGGCC AGGCTGTGAG ACAGAGGAAGC CCCTACCTT CCTCACCCCA
8701 TCCCCCTAAC TGGCAGCCAG TGGGACAGGA AGTCAGTTGT GAATCCATCC
8751 CATCCCCCGT ATGTGGCTT TCCTCTCTT CTACTGCTCT AATAATTCCC
8801 CCTAAGGAGG CAGGGAGGT GGATTCAAGG TCCCCAGAGA AAAGGGAGAC
8851 TTGAGAGAGA CGCTGCCCT GGCCCCCACCT TAGGGCCAAT CCCCATTCTC
8901 CACTCTGGG TTGCAAGGT GTGGGGCCTA CAGCTATCTC CGGGGGCTGT
8951 AAGATGCCA AAAAGTATCA CATGCAGATC GTGGGAGAAA TCCAACGCCG
9001 GTGAGTCCAG GTGGCCCAGA AGCCTGGCC ACCCGCACCT CATCCCCCAC
9051 TAAGCCTGA GCTCGGAGAG GGAGACAAGA TGAACCTTAT GAAAGTCAG
9101 TCGAAACTGT ATGACACTGA CCATGTATGA ATTATTACTA TTACCGTTT
9151 CTGAGAAGGG CGCACAACC AGCCAATGTA GGCTATTTA TGAGAAATGA
9201 GTCTTAACTG CCACACTCCC CTATAAAATC TCATTCAACT GATGCTGTTA
9251 AACAAAGCCT CTCTGAACAG CCGCTTGCTG GCTCTTGCC TTGCTCTAAT
9301 GCATTGGTTC TTTGTCCATG TAGAAAGGA ACTATTAGGT TCAACCAGAT
9351 TCATGAAGCA TCCACTCTGT GCCAGGCACC ATGCTGGGCC CTGGGAGGAG
9401 AGGGGTGACG CTTGCTCTGC AGGGTTGGAA CAGGCAAGGG AGGGAAGACC
9451 ACATAGCACC AAAGGCTCTAG GGGTCTGTGG ACTCGTGAGC ATACAGGGTT
9501 CAGAATCTGG GAGTTAACAA ACGAGGCCCT ACCACATACT GGCCCCGGGA
9551 CCTGGGCCAA GTTAGGTTCT CTCAGCTCA GTTTCCTCTT TTGAAAACA
9601 GGAGTGATGG TCCCTACCCCT ATGGGGTGGT GCTGAGGATT CAGACTGGAT
9651 GGGATAACTT AGGCAAAGAT CCCGGCACAC CATGGGGGCC TGGCTGGTCC
9701 CTGTGGGCTG GTGAAGGACT TGGCTGCCCT CCCCCACTCAC ACCCTTGGGT
9751 TCTGCCTCTT CCCTGGCTCC TCGGCAGGT CCCCACCCCG GTGTCGCTG
9801 TGGCTCACA GTGGAAGGAC ATGATAGGCA CAGCCTCTC CCTAGCCATC
9851 GTGAGCTACG TCATCAACCT GGCTATGGC CGGACCCCTGG CCAACAAGCA
9901 CGGCTACGAC GTGGATTGCA ACCAGGTAGC TCTGGCCACC CCCGGCAGGA
9951 CTGGGCAGGA CAGGTCAACT CAGGCCTGGC ATGACATATC TTGGGTGGGG
10001 AGATCATTGG GCTGAGGTTGA GGCAGGCTGC CTCGAGTGTG GGGGATAGGG
10051 GGTCTCTGA CCCTAAGAGG CTGACCTCTT CTTGACTGGG AATGTTGAC
10101 TTTATAGCCA CTGGGTCACT CTCAGGTCTT AGGCCACAG TCCAGCTTGC
10151 ATGCCTGACT GCACTTGGTC CCCGTCCCCC CCAGCCCCAC ACTGGCTTCT

FIGURE 3C

10201 AATCCTGTCC CCTCCCTGCA GGAGATGATC GCTCTCGGCT GCAGCAACTT
10251 CTTGGCTCC TTCTTTAAAAA TTCATGTCAT TTGCTGTGCG CTTTCTGTCA
10301 CTCTGGCTGT GGATGGAGCT GGAGGAAAAT CCCAGGTGAG CCTTGTCTA
10351 GGGGAGTTGG GGGGAGGTGG TAAGAGAACAA GTTGCCTCCAA AAAAGCCTGG
10401 GCACTGCAAG CCAGGCCAGC TCTTCTCCGA CCCCTTCTTC CCGTACTTAG
10451 TCTCCACTCC ACCAAAGCCA TGGATTGGAA ATAATCAAG AGCAAAAATT
10501 TCACACCTTC CCTCTATCCC CAACTCTTC TCGGAATAGG TGGCCAGCCT
10551 GTGTGTGTCT CTGGTGGTGA TGATCACCAT GCTGGTCTG GGGATCTATC
10601 TGTATCCTCT CCCTAAGGTA AGAGCCCAGC CATCGAGCAG AAGTCAACGA
10651 AAGACTCCAA TAAGAACATA CCCTGAGAGT TGTGTGGCAC TTTACGGACC
10701 ACAAAAGTGC ACTTGTGTCA TACTTAGTCT CAACACACAA CTGTGAGGTA
10751 GACAATGCAG GTTTTATCCT CCCCCATTAA CAGGTGAAGG AACTGAGTC
10801 TGAGAGTCTA AGTAACCTTG TCCATAGTGA GGCAGCTTAC AGCGCAGGGC
10851 TGGTCCAAA CTCCAGCCTT CTGGCCTAG AGTCTAATCC CTAGGCAACA
10901 TTTGCACCTA CCCACGAGTA CAAGGCTTT ATATAGCCCA GCTAGGAGGG
10951 CTCTAGGCAT GCGTCATTTA GAGATGAGGG AAGAGAGATA GGGAAAGGAT
11001 GGGGCCAGGA AGGAGCCCCAT GGCTCTAACG CCAGCACTTT CCAAACCTAA
11051 GGTGAAATGC AGAGATTTGG AGGATCAGCC AGGGGAGGTG TTCCAGAACT
11101 CCGTCTCTGT CCTGCCAGGC CTTGGGTCG GGTATGCGCA GGAGGGCAA
11151 AAGAAGGGGA GACCCCTGGGG TCCTGGAGCA ATGTTCTGCT TCTCTAGTCT
11201 GTGCTAGGAG CCTGTGATCGC TGTCAATCTC AAGAACTCCC TCAAGCAACT
11251 CACCGACCCC TACTACCTGT GGAGGAAGAG CAAGCTGGAC TGTGTAAGTA
11301 TCGGCCAGCC TCTGGGTACT GGCCATGCC CTGCCCTCTC CTCCAACCCC
11351 ACAGCCCTGT CAGCCCTGTC CTAACAAATGA ACCCTCTAGT CTGCTGCTTC
11401 CTAATTAGCA TGAGATGAGT GTTAAAGT CCGAGTTTCG AAGTGAACAA
11451 TCCTATGTT AAACCCCTAAC TCAGCCATCT GCTGGCTCCA TGGCAATAG
11501 CAAGCCCCCTT AACCTTTCCC AGTCTGGTG TCTTAACTGG GCAAATGGTT
11551 ATTTTATGCT CTCTGCCCTCC CAGGGTTTC TATGAAGAAG AAGCAAGGTA
11601 ATACAAGTAA ACATGTTGTC TACATCGTAT TTTTAACTCA ATAAAGCTTA
11651 GCTATGACTA CTTTATGACA TACAGCTTA AAAAACAAAAA GGAAATAGTT
11701 TGTTTATTA AAAAAAACCT AGAACATAAA GCCAGAGGAC CAAATCTT
11751 AGCAAGTTAC TAGACTTCCC TGGGGTTCTA TTTCTCTATC TGTAAATGGG
11801 GGTGAGACTC ATGCACTCAT GGTTGCGTC AACGCTGGTT CCGAGGATTA
11851 AATGAGATCC CAGTGGGAAA ACACCGCATG AGCGAAACAA TTCTGCAAAC
11901 ATGACTTATT GTCTGTGATTA GTCAACACACT CCACCGCATC ATCCGCTGG
11951 CATACTAATG AAGGCCAGTG TGTGGTACG ACTACTGCCCTT CTCTCCATT
12001 AAGCCCCACC ATAACCTATG GGAGAGGATT TACTAAACTT TCTTAACGGT
12051 GATGAAACCA AGGCTCAGAA TGTTAAAGTA AATTGTCAAA GGCCACAGAG
12101 GTAGGGAGTG GTAGAGTCTG GATTTAAACTT CCAAGTCTG GACTCCAGAC
12151 CTCTAGGCTG TACTGTCTCA TAGGGAAAGGC AGTCTCACCC ACCTAGGGCA
12201 GAGAAGAAAA TCCTTAAAGC CAGAGAACAGT AGTGGCTCAT CTGTGGTCAC
12251 CCAGAGAGAC AGTGTATGAGG ACAGGGAGAA AAATTATACC TCAGTTCCA
12301 GCCCAAGGAT CTGCTTGTAC CATAACCCAA CAAGCCCCCG CTATGGTGGT
12351 ATTTCTCTAG GTTCATATGG CGGCTTTGT TTCCATTGTA TCTTCACAGC
12401 AATTCTCTAC AGGAATCTGG CGACATTTAT TTCTTTAGA GGAATTCCA
12451 GGTCTTAAAAA TCTATAGGGG GCAACTATCA AAACCTTCACC CAATGTTGCC
12501 CCCTACCCAC ACACAAAACC AGGCCCCAG CCGATCAGAA AGCACTGCTG
12551 AGCTCCTGTC AGGGCCCAGC CAGCTCGCTG TGAGACAGAG AGAGGAAACT
12601 CACATTCTATT GATCACCTAC TGAGCATCCA TCACTAGGCT AGGACCGTCA
12651 CATTCTCTAA CTTTGTGATC CTTTCATGAG GTAGGCATTA TTATTCTCT
12701 TTTGTTTAC ATAGCCATTAA AAGAACAAAA TTGGGGGCTG GGTGTGCTGA
12751 CTCACACCTG TGATCTAGCA CTTTAGGGGG CTGAGGCGAGG AGGATCGCTT
12801 GAAGTCAGGA TTTCAAGGTC AGCTGGGCA GCTTAGCGAG AGCCGCTCT
12851 AGAAAAATAT AAAAGTTAGC TGGGTGTGGT GGCACGTGCC TATAGTCCCTA
12901 ACTATTCAAGG AAGGTTAGGC GGGAGCACAA CTTGGGTTCC AGGGTTTGAG
12951 GCTCCAGTGA GCTGATCTTG CCACTGCACT ACAGCCTGAG CAACAGAGCA
13001 AGACCCCTGT ACTCCAAAAAA CAAACAAACAA AACACATTTT GAAACCCAAAC
13051 AGATCTGACC CAAGATGCAT GCTCTTATAG ATGCCACCTC CCTGTGTGCT
13101 GGGGCTTCTA CTAACAAACAC AGACAAGATC AGGCAACCAC AGTCAATCTA
13151 AGGGAAAGAG GAAAGTGTAA CCAAAGCACA AATACATAAA ATATTGCAA
13201 AATGCTATTT AAAGAAAAAA AAGAGAACAGG AGGCTCTGAG GTTGTACTAA
13251 CAGAGAATGG CCTTGGCTAA TCCAGGAAGA CTTCTGTAAA GAGGTTGTTT
13301 TTTCCCCAGG TCTGCTTTTG ACATCTCTCT TTTCACAGTG CATCTGGTAA
13351 GTGAGCTTCC TCTCCTCCCT CTTCCCTCAGC CTGCCCTATG GTGTGGCAGT
13401 GGGTGTGCCC TTCTCCGTC TGGTCGTGGT CTGGCAGACT CAGTTGTAAG
13451 TGATAGCTTC CGCCCTCTTA GGCCCCACAGT CGGTTCCCTG GGCCAGCCG
13501 CAAAGGGCTT CCATGCCAGC GCCTGGCTTA GTCCACTGTAA CCTTCCACCT
13551 CTGGCCCTGG CACTGGAGGT GCTGCCAGGC CCAAAGAGAG CCCAACCCAG

FIGURE 3D

13601 CCAGGACTGT GGGCACAGTC TGGGCTGTTT GACTTCCAT ATCTTGAAAA
13651 CCCCAGAGAA AGCCAGCATA CTCTTGTGG GGATGGCTGG GGAGAGGGCA
13701 GTGGCAGAGA AAGGGGGCA AGGGCAGGTG GTGAGATTCA ACATCCTTCC
13751 AAAGACATTG CCAGAACCCC AAACCAAATG GGACCCCCACC CCAGGAGAGC
13801 GCCAGGGTGG AAGACAGAAAG CTGTGTTCTA CACACTGGGA GTATTACAGA
13851 GAAGGGGTCT TGGCCAAGGC AGGGAGTACG CTGAATGTTG GGGGAATCCT
13901 ATCTTCTCTT CTTGAGAACT CAGAACAAAGG AAATGATGAC TTCAGGGCA
13951 CTCCACACAC TTCTCCCAC CACTCTCTCC CCTGCCCTGT GGTCTGGGAG
14001 CTATGTCAG GACCTGCCTG TCATCCTCAT AGTTATAGGA GGCCACAGGC
14051 CACCAAGCAT GTGTCTCCAT TGCAAAAAGA CAGACACAGC AAGTCTGGGG
14101 GTGAGGACAG GACCCCCATCC TACCTTGGCT CTGCCCCCGC CCCAGCAGGG
14151 GCACCCCTCC AGGCCCCATGT GCCATTAGCA TTCTCTTATG TTTTCTCTT
14201 CCTGCTTCAT CCAGTCGAAA TGGCTATGCA CTGGCCCAAGG TCATGGACAC
14251 TGACATTTAT GTGAATCCCA AGACCTATAA TAGGGTAGGT AATTCAAGCT
14301 TATGACCTCC TTCTTTTGT CTGACCACCC CCAAGAAGAG GTTGCTTTT
14351 AAAGCCAATA AAGACATTT TGCAACTTGA GCTCAGTCTC CCTGTCACAG
14401 GCCCAGGATA TCCAGGGGAT TAAATACATC ACGTAATGCT CCCCCTCTCA
14451 CTTGCCCAAG TCAGAGATCT TCAGGCAAAA GGTCTATGCC AAGGTAAGGC
14501 TCAGTCCCTG GCGACCAGAG GCTCTGGACA GAGAGTGGCC GGAAAATGGA
14551 AGCAGAAGGG CGGTGGGAGC TGAGAATAGG CCACTCCAT AGAGGGTGG
14601 GGTCAAGATT GCTGTTGGCT CTCTCCCTGC AGACAGGCAT GGACCCCCAG
14651 AAAGTATTAC TAGCCAAGCA AAAATACCTC AAGAAGCAGG AGAAGCCGAG
14701 AATGAGGCC ACACAAACAGA GGAGGTCTCT ATTCTATGAAA ACCAAGGTGA
14751 ATGAAGGCCA GAAGCAGCCC CGTGCCTCTGC TCTCTGCCC ATTCTGATAC
14801 TGCCCCCTGT TACTCATGGT ACCCTGGGGG CCCCCTGTTCC CACCTGACA
14851 GGCAAGACA GAAAGTCTCT GGGAACACTG CCTGGTGGCC GCTGGCATT
14901 TTTCTTCTTT TTTTCTTTT TCTTTTTAGA GATGGAATTT TGCTCTTGT
14951 ACCCAGGCTT GAGTGAATG GCGTTATCTT GGCTCACTGC AACCTCCACC
15001 TCTGGGGTTC AAGCGATTCT CCTGCCTTAG CCTCCCAAGT CGCTGAGATT
15051 ACAGGTGCCA CCACACCCAG CTAATTTTG TATTTTTAGT AGATATTGGG
15101 TTTCACCATG TTGGCCAGGC TTGGTGTCAAAC CTCTTGACCT CAGGTGATCC
15151 ACCTACCTTA GCCTTCAAA GTGCTGGGAT TACAAGCCTG AGCCACTGCG
15201 CCCAGCCTGG GCATTTTCT TCTTGATGA GGTGCTACCA TCTCCAGGG
15251 AAGCCACTGA ACCCCCCAAGG CCCTTCTCCA TTTTCTGGCT AAGATAGGAC
15301 ATGGGCCATG GACTTTTGA CAACCCAGAG GGGGAACAGC AGTGAATTTC
15351 CTGGGAACCC CAGGCAGCCC AGGGCTAGCA AGGCTGGGGT GGCCATGGCA
15401 GTAATCCTTG TAATCCCGAC ACTTTAGGAG GCGCAGATGG GAGAATCACT
15451 CTATGAGTT CAGGAGTTCG AGACAGCCT GCCAACAGTG GCGAAACGCT
15501 GTCTCTACTA AAAATACACA AAAATTAGCC AGGCCTGGTG GTGGGCACCT
15551 GTAATCCCAG CTACTCAGGA GGCTGAGGCA CGAGAACATC TTGAACCCGG
15601 GAGGCAGAGG TTGCACTGAG CCGAGATAGT GCCACTGCAC TCCAGCCTAG
15651 GCAACAGAGG GAGACTCTGT CTCAAGAAAT AAAGGAGCTC AGTGTCCCCG
15701 GAGGGGCTTT CTCCCAGAGA GAGTGGGCTT GAGGCTTCAG TGCTCTCTT
15751 GGCTGGGTCT TCTGACTTTG TCTGGGGTTG AGGAGACCAA GTTGCAGGC
15801 CCTGCTTAAG AAAGGGCTTT GGGAGAGGCC TCTCTGGTGG AGCTTCAGG
15851 GTCTGTGTT ACCATCACCG AGGCAGTTA TTCCCTCTACA CCTACACCC
15901 CCATGCCCT GCTTCAGTCA CAGCAAGTC TGCTCAGTC TGGTGGTCCC
15951 TGAECTTGCC CACTGTCCCC ACCCTTCCAG ACTGTCCTCC TGCAAGGAGCT
16001 GCAGCAGGAC TTTGAGAAATG CGCCCCCCCAC CGACCCCCAAC ACAACCCAGA
16051 CCCCGGCTAA CGGCACCCAGC GTGTCTTATA TCACCTTCAG CCCTGACAGC
16101 TCCTCACCTG CCCAGAGTGA GCCACCCAGC TCCGCTGAGG CCCCCGGCGA
16151 GCCCACTGAC ATGCTGGCA CGCTCCCACT CTTCTGTCACC TTCCACACCC
16201 TCATCCTGGA CATGAGTGGA GTCACTTCTG TGACTTGTAT GGGCATCAAG
16251 GCCCTGGCCA AGGTGAGGCC CTCGGGGACA GCAAGCACCA CCCACTCCAC
16301 CCCCTCCGCT CTGCTCTCCA CATTCCCTT CCTGGGAGCC CTCATTTCTAG
16351 GAAGCTGAGG GAGGAAGCTC ACTGGGGAGA CTAAACAGCTC CTAGGAATCC
16401 CTCTTCTCCC CAGAGCCAC CAGGGTGGAG CATTCTCCAC AGAGCAGGCC
16451 CAGACGGGCC ATGACAATGA GTGGGGGGAC AAGTCTACCA GAGTTTCAGG
16501 CCCCTGTGCT CCCAACACCC CCAGCAGTGG CCATCCCAAG TCCCTCTAG
16551 CCATCAGGAA CCCACCCAGG TTCTCTGAGG AGGGTCCAGT TTGGCTCTG
16601 GTTCATGATC TGCTGCCCTT GTCCCTCATT CACCAGCCAC CCTAGGACAG
16651 GAGAAGAAAT AATACCACTG CCCCCACACCA TCAGGCCAAA CAGAGAGCCC
16701 ACGGGACACC TTGAATGAAT GTATCCATCT GATAACTTT CAGCAGGCCAC
16751 CGCCAATGGC GGGAGTCAGC AAACCTCAGA GCTGGCTCAG ATAGAGGCAA
16801 GCCAGGGGAA CAATGGGCAC AGAGAGTGT CGGACTGCCCT TCACCATCAA
16851 CCAGGCAGG GGCAGGGCCC ATACCCAGCC TTGGGGCTCA GCCGGCTTCC
16901 TTAGCCAGGA TCTGGAGTCC AGGCCAGCT TGCTGAAGC TCTAGACTCC
16951 CTGAGCCTCC ATCCTCCCCC GCAGCTTCTG TCTGAAGCCA CAAAGAAGTC

FIGURE 3E

17001 TGAGAATCTA AGCTACTGAA AGAAAAGATC AGCCGGGCGT GGTGGCTCAC
17051 TCCTGTAATC CCAGCACTT GGGAGGCCAA GGCAGGTGGA TCACAAGGTC
17101 AGGAGTTCAA GACCAGCTG GCAAACATGG TGAACCCCG CCTCTACTAA
17151 AAATACAAAA ATTAGCAGG TGTGGTGACG GGGCCCTGTA GTCCCAGCTA
17201 CTCGGTAGGC TGAGGGCAGAG AATTGCTGAA ACCCAGGAGG CGGAGGTTGC
17251 AGTGAGCCAA GATCGCGCCA CTGCACTCCA GCCTGGGCAA CAGAGTGAAGA
17301 CTCCATCTCA AAAGAAAAAA AAAGAAAATA TCTAGCCCCA CAAGAAGGGG
17351 CCATGGTGAC TTTAAGTGCC CGCCACGTTG GCAAAAGTCC ATTTCCGCTC
17401 CACTTCCCAG AGAAACCGTC AGCCAACACT CCAGGGAGAA GTGGTGTGCT
17451 TTGCTGCTAT TTTTGTCTT GGCTGCTGGG CTCTCAGGGT TGCTTATTG
17501 TTTGGCTTCC CCTCTGAAGT ACGTTTGTG AATCACTTT GAGACCCACT
17551 CAGAACATTC CTTTCCTTTT GCCTCCCTAC CCAAACAACA CTTCTAGCTG
17601 AGCTCCACCT ATGGGAAGAT CGGCGTGAAG GTCTTCTTGG TGAACATCCA
17651 TGGTAAGAGA AAGAGGACAT TTAGGGACTG AAAGACTGGC AAGGAGTGTG
17701 GGGTAGGAAC AGGTGGTGG GGTCTGAATA GTGAGGAGGT TGGAAACGAG
17751 AGCACCCAGC TATCCCCCAC AAGCTGCTGC CTGCTCATAA AAGCTTCAGG
17801 TACAAGTCCA AAGAGACTGG TCAGATTGCA TAAACATCCT AGGGGCCTTA
17851 GTGACAGAGT GGGGGTGAGG AGGTCTGGA GTTACAGAAG GACAGCTAGG
17901 ATTCTAATCT ACCCCATAAC TAATTGCCA CGTATCCTT GCGGAGTCAC
17951 TTTATCTCTC AAGGGATCTA TTTCTACCTA TGAAAACGA GAGGGTTGAC
18001 TAGATGGATT TGGGGATCCT CTCCCAATCA GAAACTCTGT GAATCGATAT
18051 AGGCATAGAG CACACGGTAC CCTAACATCC CAGGGAACAT ATAAATATGC
18101 AGTTTGTAG GCATACAGCC TCCAAAGGGT GCATATAACAC AGCCTCAAGG
18151 ACGTGGCCAC AGGGCGAGCAG ACATTTACAT GACTAGCATG TAGCCAAAGT
18201 GCAGAGATGT GGGAGCAAGT GCACACAGAC ACACAGGAGA ATGTGAAGGG
18251 GCACATACAC ACACACCCAG CTCCCTGCAC TGGGTAGAC CCCCCTCAGC
18301 AGGGCTGCAG TTCCCAAGCT CCGCATGGCC ACGTTCGGGG AGAGAATCTG
18351 CAGTGGCAAT GACCTGCTAT GATATGTTCT GGAGTTAGAA GCAGTGGATT
18401 CTCCCCAACC TCACTGGACA CCCCTTAGG AAACCATCTC TAGGATTAAG
18451 AGTAATCCAC ACAAACTTCC AATGCCACAC ATTGGAAGTT GCTGGAAGG
18501 TCTGGAAAAA CAAGAGGAAG GATGGGCTCT TGGGGGATAG AACTGGCAGC
18551 GGCCTCTCA AGGATGGCTT AGGCTTTCC ACTCGAATCA CCACAAAGTA
18601 CTGACTCCCT AAATCAAAC TCTTCTTCT GCTCTGGGTT GAAACTTCAG
18651 CATCCTCAAG TTTCATGTTGC CCTCTGCCGT CCAGAACTGA TATTGCACTG
18701 CCAATGCCAT GGCCTCTCAGA TACAGCAAGA GCTGGGACCT CAGGCCTCTC
18751 CCATCCCTGC TCTGGTCTCA CTATCTCCC CACCCCCAGC TCCAATCCAC
18801 AATGCTGTT ATCTTTCTGA AGGTGATCTT TTCTCTTCT AGCCCCAGGTG
18851 TACAATGACA TTAGCCATGG AGGCCTCTT GAGGATGGGA GTCTAGAATG
18901 CAAGCACGTC TTTCCCAGCA TACATGACGC AGTCCTCTT GCCCAGGCAA
18951 ATGCTAGAGA CGTGAACCCA GGACACAAC TCCAAGGGGT AAGGTTCTG
19001 CACCTGGGGG ATCCTAGGCT CCAAGGCACT GAAATAGCAG GACCAAGGAG
19051 CATTATTAGA AGAAACACAG GAGAAGGTTT AAGTTCCAAT ATCAAGTCTG
19101 CCATTTCAGT TTTCATGAACT TGTTCTTCA TCTATAGAAT GAGCACCAC
19151 AACTAACATT ACCTACCTCT CTGCATTTT CTTTTATTTT GTTTTAGGGT
19201 TAAATGATAA TTACATCTTT TGTGTCACTT GAAAGCACTT TGTGTATTGT
19251 AAAAATTCTT TATCAATATA AGTTTCTGG TTGCACAAAC ACCCAAAGCA
19301 TAGTAGAGCA GGCCCACCTC GCTGGCATCG TCCCCTGCCT CCTCCTCATC
19351 TCTTTCTAAA GGGGGCTTTC GGGAAAGGGAG GGGAGGGGAG TAAGCCTACC
19401 CATTTTAACT TACCGGAGCT TAGAGATTC AGGCTGGTGA GGGATAAAGA
19451 GATGGGGTCT GAGTTTTGTC TCAGCTTTT GACATTTAAT TTACTAGCTC
19501 AGTAAGTCAT ACAAAATGGGA TACAAATAAC ACCATCTAAA ACTCCAGAAG
19551 ACTGGGGAGT CAGAAAAATC CTACCTCTT GGGGTCCCTG CCCAGATCCC
19601 CAGTCATCTC TAGCCCTCA GGTCCCCCTCC CAGCTCAGCT CCTGCCCTG
19651 GCCTCCAAG ACTCTTGTG TGCCCCAGCC CTGGGTAAAA ACCTCCCTG
19701 CCCTCTGTGG GTCTATAAGAA AGGCTTTCT GGGCCTAGAG CAATGATTG
19751 CTCTTTGCCT TAAGAGACTG ATGAAGGTTGA AACCATCTGT TCTAAGTGT
19801 GAAAGACTGC CCAGGAACAC ACAGGGCGCT GGCTCTGCC CTCCATGCCT
19851 AGAGGGAAAC CCTGGGGAAA CAACGGGCTT TCCTGCTTCG TGAAATTG
19901 CCGCAGAGCA AAGAGGGAGA TTCTGGAGGA AGTCGATTA GTTGTATTG
19951 CCCTTAATCAT GTTCAGCTAC TCTAGTTGGT ATGTACTT GATTAGTCAT
20001 AGCACTTATA ATAATTTTAT ATTATATACTT ATATATACTT ACATATTATA
20051 GACCATTCA AGATACAAAT CACACACATA AACACACACC TTTTCAACAG
20101 CATTGTGAGG GACAAAGCAG GCAAAGTGGAG GCTGGTTATC AGACTTTAAC
20151 AGATTAGAAA ATATATTCCC AGGAGGACAG GAATTCCTTCA AGGTCAAGGCA
20201 GCTAGCCAAT AGTTTTCTA AGCTGAGTAA AACCTTCCCT GCCTCTAACG
20251 GCCCACAAAG GAGGGAAAGAC CGCGATACAC ACCTGTCTGG TATAAGGGGG
20301 AAGACCACAG CCGTGTGTT TTTGTGAGGC AGGTAAGGGG AGGGGCAAGA
20351 GGATAAGTCA TGTGTCAGGA AGCAGCGTCC AACAGAGGCC GGCCACCTGT

FIGURE 3F

20401 CCCTTTCTT GCCACCATGC ACCAACTTTG CTGTTCAGTC ACTGAAGCTC
20451 ATTCTGCACT GGCTTCTTC CTTCCAGGCT CCAGGGGATG CTGAGCTCTC
20501 CTTGTACGAC TCAGAGGAGG ACATTCGAG CTACTGGAC TTAGAGCAGG
20551 TGAGCTGAGG GAAGGGGCTG TGAGGGTGGG AGCAGGGCGA AGAGGGGAAG
20601 GATGGGGTCG CTGTCAAATA CAAGCGTTC ACTCAGCTGT CTCACCTCA
20651 GCCCAGAGCA GTCACATTCA AGGCCACAAA GATTGTGGT CATCTTGT
20701 TTTTTCTT TCCTTTCTT TTTTTTTT TTTTAATTG AGACAAAGTC
20751 TCACTCTATC ACCCAGACTG GAATGCAGTG GCATGATCTC AGCTCACTGC
20801 AACCTCTGCC CCCGGGTTT CAGAGGTTCT CCTGCCCTAG CCTCCCGAGT
20851 AGCTGGACT TCAGGGCTGC GCCCAGCTAA TTTTGTATT TTTAGTAGAG
20901 ACAGCTTTTC ACCATGTTGG CTGGGCTGGT CTCGAACCTC CGATCTCAAG
20951 CAATCTGCCT GCCTCGGTCT CCTAAGTGCC TGGAATTACAG GCATAAGCCA
21001 CGATGCCTGG CTTTGTTTT CATTCTCTC ACTCCCTGAA AGGCATCGT
21051 GGGAGAGGGT GAGTCACTGG ACCAAGTCCT AGAGAACAG TATCTATTCT
21101 TATTCTCAA CACATCACCC ACGTGACCC GAGCAAGCCA CATACACCC
21151 GGGCCCTAGT TTTTATCATC TGTGAAATTG GGGGAACAT AGGTAATACC
21201 TGTCCCATCC ACCACAAAG ATTGGCAGGG CAGTCACTTG TTCTTCTATT
21251 AATTCAGCAG GTATTTATGG CGTACCTACT GTTTGCCTGA CACAGTTCA
21301 GATGGGCACA TAGCAGTGA CAAAACAAAG GCCTCTGCCT TTTAGAAACT
21351 TACGTTATGG TAGAATAGAT GGATTTNNNN NNNNNNNNNNN NNNNNNNNN
21401 NNNNNNNNNN NNNNNNNNNNN NNNNNNNNGTCT ACAAAATGAAT TATTATTGCA
21451 TGTGACAAG CTTAAGAAC TAAAAAAAT GTGGCTGGT GCAATGGTC
21501 ACACCTGTAA TCCCAGCACT TTGGGAGGT GAGGTGGGG GACCACCTGA
21551 GGTCAAGGAGT TTGAGACCA CTTGGCCAACT ATGGCGAAAC CCCGTCTCTA
21601 CTAAAAGCAC AAAAATTAGC CAGGCCAGT GGTGCATGCC TGTAGTCCA
21651 GCTACTCGGA AGTCTGAGGC ATGAGAAATCA CTTGAACCTG GGAGGCAGAT
21701 GTTGCAGTGA GCCGAGATCG TGCCACTGCA CTCCAGCTTG GGTGACAGAG
21751 CTAGACTGTC CAAAACAAAC ACAAAACAAAC CAAAACCTAA AAGATATGTG
21801 GATATGAGGG ATCACCATCC CCATAGGGCC CCTGGATTAA CACCAACCA
21851 CCAATGCCCT GAATTAAAG AAACAGATG ACTAGGTTTG GAGAAATCTG
21901 GCTTTGGGTC TATGAGAAGT AGTGTCTCTC TTTGTGCCTC TTCCCATTCT
21951 TTTTGACATT GAGCTCATG TGCTCTGAA TCCGTCTCTC ACAGTGTGA
22001 TGGCAGGTGG GACAGATTAG AAAATAGAGC TGAGGCCACA GAGATTTGGC
22051 AGACTGATT CCGTGCCCT TTGGAATCTC CAGCACATT CAAAAGCCT
22101 GGATAGGACC AAAAATAGCT ATCAACGTGA GAAAGGACTT CAGAGCTTGT
22151 CTACTGCCAA CCCTCATTTT ACCCAATGAG GAAAGTGAAG CTATTAGGG
22201 GCGAGGGACA CGTGGAAAGGT CACACAGCAC ACAGGAGGTG ATTACATGT
22251 AGATTTAGC ACCTGCTCTT GCCACGCTGG ACTGGTTCAC CTCTTAGGCT
22301 GACCTGCCT CTCCCCCTGTT CACACACACT CTCGCACACA CACACACACA
22351 CACACACACA CACAGGTGCT TTGTTCTGGC CAGGGGTTCC TAGGGTCACC
22401 TCTTGGTTCG AGCCACTGTG ACCCCAACGT GTCTAACCTC TCTCTTCCCC
22451 TCCCACCTTC TTCTCTGTT CTCGAGGA GATGTTGGGG AGCATGTTTC
22501 ACGCAGAGAC CCTGACCGCC CTGTGAGGGC TCAGCCAGTC CTCTGCTGC
22551 CTACAGAGTG CCTGGCACTT GGGACTTCCA TAAAGGATGA GCCTGGGTC
22601 ACAGGGGTG TCGGGCGGAG GAAAGTGCAT CCCCCAGAGC TTGGGTTCC
22651 CTCTCTCTC CCCCTCTCTC CTCCCTTCTC TCCCTCCCCG CATCTCCAGA
22701 GAGAGCCTCT CAGCAGCAGG GGGGTGCTAC CCTTACAGGA GTGAGAGTCT
22751 GGTGAGCCCA CTCTTACCC GTCAGGCCCT GGGCGCAATG GACAAGCCTC
22801 CTGCTCACT CACCCCCACCC ACCTCTGCC TGTCTTGGC AGCTGAAGGA
22851 CACCTTGACT TCCAGCTTTT ACGAGTGAAC CAAAAACAGA AGGACAAGTA
22901 CAACTGTGCT GCCCTGCTGT ACAAGCTTC AAAAGTGTCC CAGAGCCAC
22951 ACGGCTCGGT GTCAGATGGT GTCAGGCTGT CACGGACATA GGGATAAACT
23001 TGGTAGGAC TCTGGTTGCT CTTCCCCAGC TGCTCAACT CTGTCTCTGG
23051 CAGCTCTGCA CCCAGGGGAC ATGTGCTCTC CACACCCAGG AGTCTAGGCC
23101 TTGGTAACTA TGCGCCCCCCC GTCCATCATC CCAAGGCTG CCAAAACAC
23151 CACTGCTGTC AGCAAGCACA TCAGACTCTA GCCTGGACAG TGGCCAGGAC
23201 CGTCAGAGACC ACCAGAGCTA CCTCCCCGGG GACAGCCAC TAAGGTTCTG
23251 CCTCAGCCTC CTGAAACATC ACTGCCCTCA GAGGCTGCTC CTTCCCCCTG
23301 GAGGCTGGCT AGAAACCCCA AAGAGGGGG TGGGTAGCTG GCAGAACAT
23351 CTGGCATCTC AGTAATAGAT ACCAGTTATT CTGCACAAAA CTTTGGGAA
23401 TCCCTTTG CACCCAGAGA CTCAGAGGGG AAGAGGGTGC TAGTACCAAC
23451 ACAGGGAAAA CGGATGGGAC CTGGGGCCAG ACAGTCCCCC TTGACCCCCAG
23501 GGCCCATCAG GGAAATGCCT CCCTTGGTA AATCTGCCCT ATCCCTTCTT
23551 ACCTGGAAA GAGCAATCA TGTTAACTCT TCCTTATCAG CCTGTGGCCC
23601 AGAGACACAA TGGGGTCCTT CTGAGGCAA AGGTGGAAGT CCTCCAGGG
23651 TCCGCTACAT CCCCTAACTG CATGCAGATG TGAAAGGGG CTGATCCAGA
23701 TTGGGTCTTC CTGACAGGA AGACTCTTA ACACCCCTAG GACCTCAGGC
23751 CATCTCTCC TATGAAGATG AAAATAGGGG TTAAGTTTTC CATATGTACA

FIGURE 3G

23801 AGGAGGTATT GAGAGGAACC CTACTGTTGA CTTGAAAATA AATAGGTTCC
 23851 ATGTGTAAGT GTTTGTAAA ATTCAGTGG AAATGCACAG AAAATCTTCT
 23901 GGCCTCTCAT CACTGCTTTT CTCAAGCTTC TTCAGCTTAA CAACCCCTTC
 23951 CCTAACAGGT TGGGCTGGCC CAGCCTAGGA AACACATCCCC ATTTCTAAT
 24001 TCAGCCAGAC CTGCGTTGTG TGCTGTGTG TTGAGTGAGC TGGTCAGCTA
 24051 ACAAGTCTTC TTAGAGTTAA AGGAGGGGGT GCTGGCCAAG AGCCAACACA
 24101 TTCTTGGCCC AGGAGCATTG CTTTTCTGTG AATTCTATTAT GCCATCTGGC
 24151 TGCCAATGGA ACTCAAAACT TGGAAAGGCAG AGGACAATGT TATCTGGGAT
 24201 TCACCGTGCA CAGCACCCGA AGTGCCTAAAT TCCAGGAGGA CAAGAGCCTT
 24251 AGCCAATGAC AACTCACTCT CCCCTACTCC ACCTCCTTCC AAGTCCAGCT
 24301 CAGGCCAGG AGGTGGGAGA AGGTACAGA GCCTCAGGAA TTTCCAAGTC
 24351 AGAGTCCCCT TTGAACCAAG TATCTAGATC CCCTGAGGAC TTGATGAAGT
 24401 GATCCTTAAC CCCCCAAGTAA TCATTAACCC CCAGACCAGC CTCAGAACTG
 24451 AAGGAGATTG TTGACCCAGT GACCTGGAGT TGAGGCTCAG GGAGAGATCT
 24501 GCCACATGTC TGAGGGTTGC AGAGCC (SEQ ID NO:3)

FEATURES:

Start: 1997
 Exon: 1997-2121
 Intron: 2122-4732
 Exon: 4733-4872
 Intron: 4873-5004
 Exon: 5005-5115
 Intron: 5116-5781
 Exon: 5782-5957
 Intron: 5958-7770
 Exon: 7771-7935
 Intron: 7936-8470
 Exon: 8471-8623
 Intron: 8624-8917
 Exon: 8918-9000
 Intron: 9001-9777
 Exon: 9778-9925
 Intron: 9926-10221
 Exon: 10222-10335
 Intron: 10336-10539
 Exon: 10540-10617
 Intron: 10618-11197
 Exon: 11198-11293
 Intron: 11294-13338
 Exon: 13339-13445
 Intron: 13446-14214
 Exon: 14215-14284
 Intron: 14285-14400
 Exon: 14401-14493
 Intron: 14494-15980
 Exon: 15981-16262
 Intron: 16263-17597
 Exon: 17598-17652
 Intron: 17653-18842
 Exon: 18843-18988
 Intron: 18989-20477
 Exon: 20478-20549
 Intron: 20550-22478
 Exon: 22479-22523
 Stop: 22524

CHROMOSOME MAP POSITION:
 Chromosome 1

ALLELIC VARIANTS (SNPs):

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
48	C	G	Beyond ORF(5')			
132	G	A	Beyond ORF(5')			
724	A	C	Beyond ORF(5')			
1558	C	G	Beyond ORF(5')			

FIGURE 3H

1577	A	G	Beyond ORF(5')			
2487	C	A	Intron			
2634	T	C	Intron			
4352	A	G	Intron			
5157	A	C	Intron			
5658	A	T	Intron			
5945	T	C	Exon	180	T	T
6281	C	T	Intron			
6452	G	C	Intron			
6610	T	G	Intron			
7247	T	C	Intron			
7360	A	G	Intron			
7644	A	T	Intron			
8127	A	C	Intron			
8317	G	A	Intron			
9079	G	A	Intron			
9537	G	T	Intron			
12302	C	G	Intron			
12354	C	T	Intron			
12487	C	T	Intron			
13198	-	A	Intron			
13257	A	G	Intron			
14541	G	A	Intron			
14545	A	G	Intron			
15041	C	A	Intron			
15053	A	C	Intron			
15065	A	G	Intron			
15108	A	C	Intron			
16274	-	G	Intron			
17424	C	T	Intron			
17627	G	A	Exon	657	V	V
18427	T	C	Intron			
18813	C	G	Intron			
19035	T	C	Intron			
19182	T	C	Intron			
19508	-	G C	Intron			
19571	T	G C	Intron			
20147	T	G	Intron			
20180	G	A	Intron			
20584	A	T	Intron			
20717	T	C	Intron			
20894	A	G	Intron			
21787	-	A C	Intron			
22264	T	C	Intron			
22338	-	C A	Intron			
23363	T	C	Beyond ORF(3')			
23688	G	A	Beyond ORF(3')			
24210	A	C	Beyond ORF(3')			

Context:

DNA Position

48 CTGGGTTCTATGTGGGGAGGTATGCTCCCCACTCATTGAGCCCC
 [C, G]
 CAGGCAAACCACCTGGACAGCCAGACCCATGCAGACTCTGGAGCAGGTGGAGAGGAAGAG
 TGAGACCACCCCGCCTACGGGCGGTGAAGGGCCGGCAGCCTCTGAATAGTCTCTGCTAG
 GAGGTAGAAAGCACCCCTCCATCTTAATCATAGTAATCATGCCACTACCATTACTGGG
 TGCCTATAAAAGGCCAGCCTCTTCATACACATGATCTCACTGAATCCTCATAGCATCTGC
 CTGCGACTGTTATTATCCCCATTACAGATGAAGAAACTGAATCTTGAACCCAGGTAT

132 CTGGGTTCTATGTGGGGAGGTATGCTCCCCACTCATTGAGCCCCCCCCAGGCAAACCA
 CTGGACAGCCAGACCCATGCAGACTCTGGAGCAGGTGGAGAGGAAGAGTGAACCC
 GCCTCACGGGC
 [G, A]
 GTGAAGGGCCGGCAGCCTCTGAATAGTCTCTGCTAGGAGGTAGAAAGCACCC
 TAATCATAGTAATCATGCCACTACCATTACTGGGTGCTATAAAAGGCCAGCCTCTTC
 ATACACATGATCTCACTGAATCCTCATAGCATCTGCCTGCCACTGTTATTATCCCCATT

FIGURE 3I

ACAGATGAAGAAACTGAATCTTGAACCCAGGTATCTGGCTCTCAAACCTGTGCTTT
 TCCCTAAGCCACCCGGTCTCATTTCTCCACTGAAATGTCTACATGCCATTGCCCTT

724 ATTGCCCTTACTCATTCTGCCCATGTCCTCTCCAAAACACCATTTATCAATTGCTCAA
 CAAGTATGTGTTGAGTACACACTAAGGGCCAGGCAGGGGCTGGGCACAGGCCTGGGG
 TAGGTTATTCTCCACCTCGCTCTGCTGGGTATCACCTGTGGGTCTTGGCCGGCAT
 CCCACCCCTACCTGTAGTTCAAGTGGACCTGGGATCCAAAGACCAAATGAATGGAATGC
 ACCAGGCCAGCCTTACCAACATTGAGCACAATCTTATTCTAAATAGAAAATCACATTG
 [A, C]
 TCACACTTACATTTCACAAACCCCTTATCCATTAACTCATTGATCTTACAACA
 ACCCTGTGAGATATGTCGTTACTCCACTTTAGTGTACAGAATCTGAGGTTGAAAAG
 TAATGTCGACCATCTGCCTCATTAATAAAAGCAGGATTAACCCAGGCTCTGGACCC
 CCACAAAAGGCATTAAGCAACCTGCTCCCTCTGACAACCTCCCTGTCACCCAGGCT
 CCTCTGGGAAGTTGGGGCATCTAGCCCCAAGTAGTTACTCATTTCAACCCATCT

1558 TCAGCTCTGCCATCTCAGCTCTGGAACGTCAGCCAGGGTGCACAAAAGTGAGGAGGA
 GAGGAGCGGCAGTACACAAGGGTGGGGAAAGATTAGGCACAGGAAGCCGTGGGAGAGAG
 AGCCGGCAGGTGGACCATCTGGTTTCCCCACACACACATTGTCCTGGGAAACCTG
 TTGGTGAAGTTCTAGATGCTTATCCAAGAAGGGTCTTGTAGGTATCTCAGCTATCC
 CCTGCCTCTAGGCAAGCTGTTTCTGTTCTTCAAGCTGACTGGCTGAATGGTAGGAG
 [C, G]
 CTTCTGCAGGGAAACTAAGGCTGGGAAGGGAGTATGGCTTGTGGGACACCAAGGGT
 CAGGGGAGGGGAGGGTCCACCTGTAAGTGGGGCTCTGCCCTGTGATTCCCC
 TTGCTGGTCTAGGGGTGATGGTGACGCCACAGGTGTGGAGTGCCAGCACGTG
 CTGAGCGCAGAAAACAGCCAGGGTAGTCTATGCATCATGTCGCTGGGAAGGAAG
 GCCACTCGGAGCAGGGAGTCTGACGGAAAATGACAGAGGAAGGGAGGCACCTTGCT

1577 CTCCCTGGAACGT CAGCCAGGTTGCACAAAAGTGAGGAGGAGAGGAGCGCAGTACACAA
 GGGTGGGGAAAGATTAGGCACAGGAAGCCGTGGGAGAGAGAGCCGGCAGGTGGACCATC
 CTGGTTTCCCCACACACACATTGTCCTGGGAAACCTGTTGGTGAAGTTCTAGATGT
 CTTATCCAAGAAGGGTCTCTTGAGGTATCTCAGCTATCCCCCTGCTCTAGGCAAGCT
 GTTTCTGTTCTTCAAGCTGACTGGCTGAATGGTAGGACCTTCTGCCAGGGAAACT
 [A, G]
 AGGTCTGGGAAGGGAGTATGGCTTGTGGGACACCAGGGTCAAGGGAGGGAGGGTCCA
 CCTGCTGAATCAAGTGGGGCTCTGCCCTGTGATTCCCCTTGCTGGTGTCTAGTGG
 GGGTGTGGTACGCCACAGGTGTGGAGTGCAGCCACGTGCTGAGGCCAAGCAAAACA
 GCCAGGGTAGTCTATGCATCATCAGTGCCTGGGAAGGAAGGCCACTGCGAGCAGGGAGT
 CTGACGGAAAAACTTGACAGAGGAAGGGAGGCACCTGCTTATGGGGCGGGGAAGGC

2487 ACACGGCTCTGCACTGGTATCCCTAACATGGGTTAACGGGAGGCCCTGGGAAGTGAGG
 TTCTGAATGATGAATTAAAGATCTACAAACCTCATCTGACTGAGACCCCCAGGGAGGAT
 GGGGAGCAGGAGCAAGAACCATCCAGAACGGTTATATGGCATCTCCAAACCCCTGCATGG
 CATCTCCCATATTCTCAATTACCCGGTCTCTGGGTTTGTAAAGCATGGTAGATGA
 GCATCTACGTTATGGAGGGTGGGAGCATCAGAGCCCTACTCCATGCCCTGTTCCCTC
 [C, A]
 TTACAAAAAAACTCTGAAGTTACCATCACCCAGGTTCTTGTCTTCCCTCCGGATG
 TTCTTCCCTCACTGGTCCAGAGAATGCCAAAGGGGCCCTAAATTCTGAACCTTCC
 TGAGGGGACCTACCAGGGTAGTCTACCAGCGCCACGGGTCTTCCACTCTCATCTC
 CTGGAAATGCATGGTGGGTATGAAACCTTGTCCCTAACGGCGTACACAAGGTGATC
 CATACCCACACCCAGGAGGCTGGGCTGCGGGTGTACCCCTCCCCATTCCAGACTCCT

2634 AGGGTTATATGGCATTCCAAACCCCTGCATGGCATCTCCATATTCTAACCCGG
 GTCTCTGGGTTTGTAAAGGCATGGTAGATGAGCATCTACGTTATGGAGGGGGGGAG
 CATCAGAGCCCTACTCCATGCCCTGTCCTCCCTTACAAAAAAACCTGAAGTTACCAT
 CACCCAGGTTCTTGTCTTCCCTCCGGATGTTCTCCCTCACCTGGTCCAGAGAA
 TGCCAAAAGGAGGCCCTAAATTCTGAACCTTCTGAGGGGACCTACCAAGGGTAGTCC
 [T, C]
 ACCAGCGCCCAGGGTCTTCAACTCTCATCTCCCTGGAAATGCATGGTGGGTATGAAAC
 CTTGTCCTAACGGCGTACACAAGGTGATCCATACCCACACCCAGGAGGCTGGGC
 TGCGGGTAGTCAACCTCCCCATTCCAGACTCTGGCAGACCTCCCTGGCCAGCTATAG
 GCCAACTCACTCTCCCTCACTCCCTGGGGAAACGGCTGATTCAAGTACCTGGATTGAGG
 TCACTGGCAATGGCTGAAGTGGAGACGCAGGTGGAACTGGTTCAAGGCCGGGGAAATCACC

4352 ATTGGAGTTACACACATAAAGGAGTAGTGAGTCAGCAGAGTGACCCCTGCAGGAACAATA
 GAGCCTCTTTCAAGGAAGTCTAAGAAAAATGGCAGCAGGGCAGGGCCACTCGGGTG
 TATTCACTATTCAATTCAACAAATATTACTAACGGAGAGCTAGACTTCTGCCATGCGTGGTGGGGTTGCTGCTA
 GTGGGAGAGACAGACAAAAAGCAAGGAATGCACACACAGGATGCACACAGCGCAGGA

FIGURE 3J

[A, G]
 CCAAGGTGCGATTACCGGCCCTGGGATCAGACAGACAGGACTCAGAGGAGACTTCCCA
 GAGAAAAGGCCATCTGAGCCAAGGGATGGATCTGATAACCTCCGAAGGCTGAGCCACCTAA
 CACTCATACCTTAAGCCAAGTCTTATAAACTCCCCAGGTAAAGCAGCTGCAGTCAGAAG
 ACCTCCAGCTAATGCCAGGACAAGTTGAGCTCTCAAGAAAAGTTCCCTGCTTTC
 TTCTCAATATCCCTGGCACACAGTCAGTGAATTGAATGAACCAATGAATGAG
 5157 ATCCAGGTCCCACAAGGTGAAGGGCTCCTTCAGCCAGGGCTGGATTGCCACTCCCTCA
 CCCTTCCCTCCTCATCCCACTCCATCCCTCTGTGATCCCCATAAGCTAGTCATGCTGC
 TGAGCTTCAGTCTCGTGTCTCTGCAGGCATGGCATTGCTCTGCTGGCCAACCTTCT
 GCAGTCATGGCCTACTCTCCCTCTCCCCCTCTGACCTACTTCTCTGGGGGT
 GTTCAACAGATGGTGCAGGTAAAGGCCCTCTCCCCCTGGGCAGGCAGGATGACCCAGACC
 [A, C]
 CAAGGATGGGAGGTGGCAAAGGGGCCCTGGGAGATTTCATCTGCATTCTCTGGAG
 TTGTTCTGGTCAGTCTAGGGAAATGGTCACTGTGAATGTCATTTCCAGGTCTCGGTG
 ACCTTGGAGAAACCACTGAGCCTTTGAGTTAGTTAGCATTACCTGTTCCATCTTCT
 CCTAGGAATGAGAGGAAGACTTAGCAGAACAAAGATACCATATGCTATAACATGCTTAA
 ACAGATGTGAGAAATCACCCTAACTCCCTGGTGGTCCCAGCCGGCACTACAGGGAC
 5658 TTAGCAGAACAGATATACCATATGCTATAACATGCTTAAACAGATGTGAGAAATCACCA
 TCTAACTCCCTGGTTGGTCCAGCCGGCACTACAGGGACATTGGACTCTCTGGTGT
 AAGTGAGATGGAGGAAGCCTGGTACAAGGCTGGTTGGTCAAGGCTGGTTCTGGTCAAGGCTCTGCTTATA
 TTTCTTATTCAGTTCTACGGTGTGGTCAAAACCTCTGATTTTGCTGGCTGGTGGGGGT
 AAAAGCACCTTGAAAGCATAGATCATGGTTAGGTGAGTGAGTGGTTTATTATTGTGTTGG
 [A, T]
 GAAGAGCCTTGGAGGTGCAGGGATCCATCCCCCTGGGGCTGGGAAGCATTCTGGGCC
 TTTCTGGTTTCCATCGGTGTGGTCAAAACCTCTGATTTTGCTGGCTGGTGGGGCACCA
 CAGGTACCTTGCCTTACAGCATCTGGTGGTAAACATCTGCTGCGAGCTGGGCCAG
 AGTCGAAATTCCAGGTCTCAACAAATGCCACCAATGAGAGCTATGTGGACACAGCAGCCA
 TGGAGGCTGAGAGGCTGCACGTGTCAGCTAGCCTGCCTACTGCCATCATCCAGG
 5945 ATTATTGTGTTGGAGAAGAGCCTGGAGGTGCAGGGATCCATCCCCCTGGGGCTGGGAAG
 CATTCCGGCCCTTCTGGTTCCATCGGTGTGGTCAAACCTCTGATTTTGCTGGC
 TGGGTGGGGCACACAGGTACCTTGCCGTTATCAGCATCTGGTGGTAAACATCTGCT
 GCAGCTGGGCCAGAGTCGAAATTCCAGGTCTCAACAAATGCCACCAATGAGAGCTATGT
 GGACACAGCAGCCATGGAGGCTGAGAGGCTGCACGTGTCAGCTAGCCTGCCTCAC
 [T, C]
 GCCATCATCCAGGTGAGGGGGCAGCCCCAACCCCTGCTAGAAGGGCATCAGACCAACCTG
 CCCCTCCCTCAAAGCCTTAGCTTGATGCTAAATCTGATTAGGGGCTGGGTGGAGG
 CTCATGCCGTAAATCCAGCACTTGGGAGGCTGAGGAGGGTGGATCACTTGAGGTGAGG
 AGTTGAGACCACCTTGACCAACGTGATGAAACCCATCTCTACCAAAAATACAAAATA
 ATCCAGGCTTGGTAGTATGCCCTGAGTCCACCTACTCAGGAGGCTGAGGCAGGGAGA
 6281 GCTAGAAGGGCATCAGACCAACCTGCCCTCCCTCAAAGCCTTAGCTTGATGCTAAATC
 TGATTTAGGGGCTGGGTGTGGAGGCTCATGCCGTAAATCCAGCACTTGGGAGGCTGA
 GGAGGGTGGATCACTTGAGGTAGGGAGTTGAGACCACCTTGACCAACGTGATGAAACCC
 CATCTCTACCAAAAATACAAAATAATCCAGGCTTGGTAGTATGCCCTGAGTCCCACC
 TACTCAGGAGGCTGAGGAGGAAATCATTGAATCCGGGAGGAGAGGTTGAGTCAGTGC
 [C, T]
 GAGATCGGCCACTGCACTCCAGGCTGGGTGACAGAGCGAGACTCCGCTCTAAAAAAA
 AAAAAAAAAAAAAAAAAAAACAGTTAGGGCTCACCTCTCCCTCCCATCCAGG
 GCTAAAGTGAACCTGAAAATTAAACAGTATCTCTCATCTGATGAGCAGGACCATACA
 AAAAAACAAACAGCTGTACCTGGTAAACTGTCCTGAGCTTAAACCTGTAAGACTCAC
 AGCCTCTCCATTATCCGTGGAGAACCCAACCTCTGCCAGCATAGCTTGAGCA
 6452 ATGAAACCCATCTTACCAAAAATACAAAATAATCCAGGCTGGTAGTATGCCCTGT
 AGTCCCACCTACTCAGGAGGCTGAGGAGGAAATCACTGTAATCCGGGAGGAGAGGTT
 GCAGTGGAGCTGAGATCGGCCACTGCACTCCAGGCTGGGTGACAGAGCGAGACTCCGCT
 CAAAAAAGGGCTAAAGTGAACCTGAAAATTAAACAGTATCTCTCATCTGATGAGC
 [G, C]
 GACCATACAAAAAACAAACAGCTGTACCTGGTAAACTGTCCTGAGCTTAAACCTGTA
 AAGACTCACAGCCTCTCATTATCCGTGGAGAACCCAACCTCTGCCAGCATAGTC
 TTGCAAGCTGTAATTCTCTAACATCCCTCACTCCGCTCCAGCCTCTGCTCCAAG
 CCACAGCAGCAGTTGACAACATAAAATTGAGCTCTGCAATGGTGCAGGAGGATTCTGC
 TAGGTTTATGAAGGGAAAGCACAACATGACAGAATGCAAGAGCAAAACACAGTCCAGAG
 6610 GTGACAGAGCGAGACTCCGCTCTAAAAACCAAGTTA

FIGURE 3K

GGGCTCACCTCCTCCCTCCCCATCCCAGGGCTAAAGTGAACCTTAAAATTAAACAGT
 ATCTCCTCATCTGCATGTAGCAGGACCATACAAAAAAACACAGCTGTACCTGGTTAAC
 TGTCTGAGCTTAAACCTGTAAAAGACTCACAGCTCTCCATTATCCCAGGAGAAA
 CCAACTCTCTGCCAGCATAGCTTGAGACTGCTAATTTCTAACATCCCTCACTCC
 [T, G]
 CTCCAGCCTCCTGTCCAAGCCACAGCAGCAGTTGCACAACATAAAATTGAGCTTCTGC
 AAATGGTTGCAAAGGATTCTGTAGGTTTATGAAGGGAAAGCACAACATGACAGAAATGCA
 AGAGCAAACACAGTCCCAGAGAGCGCTTTCATTCACTCATTCACTCGGTTTGTGCC
 AAGAACTAGGCTAACCCCTGGGATACAAAGATAAGTAAGAAAGAGGTCAATTACAAGT
 TGCTCACAGCCCAGCAGAGGAAGGCCATGTCAACAGATAAATTGTATGCACTGAGAT

7247 GACACAGAGCAGAGTCACGGAGGACCTCAAAGAGGAGGTGACACTCCACCTCTTTAAAG
 GATGAGAACTTAACCAAGGAAACAAGGTATACAGAGGATGGTCAGGCAGAAGGGAAACAGTG
 CCTAAAAACACTGAGGCTGAGAGAGTGTGATCTGCGAGGCAAAGTAAGGGGTTGGTG
 TGGCTGGAGGGTAGAGGGCCAGAAGAGGATGGAAAGTAGGCAGGAGCCAGACAATGAG
 ATCTGGGTCTGTTCTGACAGCGACTTGGGCTGATTGGCAGTTATAAGGATCGTT
 [T, C]
 GGGCTACACAATGATGAGTGGGAGGTGGATTAGAATCAAGGCAGGGGACCTGTTGGGAGA
 CTCTGCAGAGGCCAGGCAGGAATAATGCAAGCGAAGACAGGTAGAGAAAGAGATGGGG
 CTGGACTTAAAAGAATGTTTACCAAGGAGCTTGGTGTAGACTGGATGTGGGAGGTAAAG
 GGAGGATGACTCTCAAGTTTGGTGGCAACAGGTAAATGATGGTGTCAATTACTGA
 GAGAGAAAACACTGGGGAGGACTAGACTTATTACAGATAAGCCAAAGGCCAGAGAGGT

7360 AACAGTGCCTAAAAACACTGAGGCTGAGAGAGTGTGATCTGCGAGGCAAAGTAAGGGG
 CTTGGTGTGGCTGGAGGGTAGAGGGCCAGAAGAGGATGGAAAAGTAGGCAGGAGCCAGA
 CAATGAGATCTGGGCTCTGTTCTGACAGCGACTTGGGCTGATTGGCAGTTATAAG
 GATCGTTGGGCTACACAATGATGAGTGGGAGGTGGATTAGAATCAAGGCAGGGACCTG
 TTGGGAGACTCTGCAGAGGCCAGGCAGGAATAATGCAAGCGAAGACAGGTAGAGAAAAG
 [A, G]
 GATGGGGCTGGACTTAAAAGAATGTTTACCAAGGAGCTTGGTGTAGACTGGATGTGGG
 AGGTAAAGGGAGGATGACTCTCAAGTTTGGGCAACAGGTAAATGATGGTGTCACT
 TTACTGAGAGAAAACACTGGGGAGGACTAGACTTATTACAGATAAGCCAAGGCCA
 GAGAGGTGTGACAGAAAGGCCATGCTAAAGGAGCTGAAGGTGTGATGGCAGCCA
 TGTAGAGCACAGTGAAGGGCAGGTGAAGGTACAGATGGTCAATTCCCTCAAGCTACTG

7644 GACCAGGTAGAGAAAGAGATGGGCTGGACTTGAAAGAATGTTTACCAAGGAGCTTGGT
 GATAGACTGGATGTGGAGGTAAAGGGAGGATGACTCTCAAGTTTGGGGCAACCGAG
 GTTAATGATGGTGTCTTACTGAGAGAAAACACTGGGGGAGGACTAGACTTATT
 CAGATAAGCCAAGCCAGAGAGGTGTGACAGAAAGGCCATGCTCAAAGGAGCTGA
 AGGTCTGATGGCAGCATGTAGACAGCACAGTGAAGGGCAGGTGAAGGTACAGATGGTCCA
 [A, T]
 TCCCTCAAGCTACTGCTACGCTAGGACTGCACGGAGCTCCAGACCTGCGTGTGTGGG
 GCAGGGCTGGACTGCTGAACACATTGGTCTTCCGCCACCAACCACCCCTTCTCC
 TCTCAGATGGTCTGGCTTCTGAGCTTGGCTTGTGCCATCTACCTCTCCGAGTCC
 TTCTCCGGGCTTCTGACGGCCGCCGCGCTGAGCTCTGATTTGGTGTCAAGTAC
 ATCTTCGGACTGACCATCCCCCTTACAGGCCAGGGTCCATGCTTTGTGAGTCTG

8127 CATCCGGGGCTTCTGACGGCCGCCGCGCTGCAGATCCTGATTCGGTGTCAAGTACAT
 CTTCGGACTGACCATCCCCCTTACACAGGCCAGGGTCCATGCTTTGTGAGTCTGGG
 GATGCACCCCTGCCATTGGAGCAAGGCTCCAGCAGACACATGAGGAGGATGACTGTTT
 AAGATGTCGTGAGCTCTCATTGCAAGGGCTGGTTAGCTGTTGAGAGGAGATTCTG
 AGGGGGTTCTGTCTGGGAGGGTCAAAGTCATGACTCACAGAGGTCTTGGTAGTTAAT
 [A, C]
 CCTGCAGAAAAGAGCTGTACATTCTCCGCCAGTTCCCATTCTAGTGCCTCAACCCCTCC
 CTGCCTGGAAAAGTCTGCCTTATGCTAACTCCATCCCCCTCCCTCAGCCCAAACCTCT
 TCTAAAGAAAAGAAAGCATTCTTCTAGACAAGTCCCCTGTCAGGGTTTAAATTCTGCTTGGTGTGCC
 GGGCGGTGGCGACGGGACAGGGTCTGATCAGGGTTTAAATTCTGCTTGGTGTGCC
 CCATTAGCTTGTGATGGCATCCCTCCCTGGGTCAAGACACCCAAAGGTGGGTATTATGGG

8317 GAGCTCCTATTGCAAGGGCTGGTTAGCTGTTGTCAGAGAGGATTCTGAGGGGGTTTC
 TGCTTGGGAGGGTCAAAGTCATGACTCACAGAGGTTCTGGTAGTTAACCTGCAAGAA
 AAGAGCTGTACATTCTCCGCCAGTTCCCATTCTAGTGCCTCAACCCCTCCCTGCC
 AAGTCCTGCCATTGCTAACTCCATCCCCCTCCCTCAGCCCAAACCTCTAAAGAA
 AAAGAAAGCATTCTTCTAGCACAAGTCCCCTGTCAGGGTTTGGGAAAGGGCGGTGG
 [G, A]
 CGACGGGACAGGGTCTGATCAGGGTTTAAATTCTGCTTGGTGTGCCATTAGCTT
 TGATGGCATCCCTCCCTGGGTCAAGACACCCAAAGGTGGGTATTATGGGAAAGAGGGT
 GGGAGCCTGTGAGCATGCTTTCCCCAGACCTTCATTGACATTGCAAAACCTC

FIGURE 3L

CCCCACACCAACATGCCCTCGCTCATCTCGCTCATCAGCGGTGCCTCCCTGGTGCTG
 GTGAAGGAGCTCAATGCTCGCTACATGACAAGATTGCTTCCCATCCCTACAGAGATG

9079 TTCTACTGCTAATAATTCCCCCTAAGGAGGCAGGGAGTGGGATTCAAGGTCCCCAGA
 GAAAAGGGAGACTTGAGAGAGACGCCCTGCCCTGGCCCCACCTTAGGGCAATCCCCATT
 TCCACTCTGGGTTTGAGGTGGTGGCAACAGTATCTCGGGGCTGTAAGATGCC
 CAAAAAGTATCACATGAGATCGTGGAGAAATCCAACGCCGTGAGTCCAGGTGGCCA
 GAAGCCTGCCACCCGACCTCATGCCCACTAAGGCTGAGCTGGAGAGGGAGACAA
 [G, A]
 ATGAACCTATGAAAGTGCAGTCGAAACTGTATGACACTGACCATGTATGAATTATTACT
 ATTACCGTTCTGAGAAGGGCCACAACCAAGCCAATGTAGGCTATTTATGAGAAATG
 AGTCTTAACGCTGCCACACTCCCCATAAATCTCATTCAACTGATGCTGTTAACAAAGCC
 TCTCTGAACAGCCGCTGGCTGGCTCTTGCTTAATGCAATTGGTTCTTGCCAT
 GTAGAAAGGAACTATTAGGTTAACCGAGATTCAAGGCATCCACTCTGTGCCAGGCAC

9537 AACTGATGCTGTTAACAAAGCCTCTGAAACAGCCGCTGGCTCTTGCCCTGCTC
 TAATGCATTGGTTCTTGTCCATGTAGAAAGGAACTATTAGGTTAACCGAGATTCA
 AGCATCCACTCTGTCAGGCACCATGCTGGCCCTGGGAGGAGAGGGGTGACGCTTGTC
 CTGCAAGGGTTGAAAGGCAAGGGAGGGAGACACATAGCACCAAGGTCTAGGGTCT
 GTGGACTCGTGAGCATAAGGTTAGAATCTGGAGTTAACAAACGAGGCCCTACCACA
 [G, T]
 ACTGGCCCCGGGACCTGGGCAAGTTAGGTTCTCTGAGCCTCAGTTCCCTTTGTA
 AAACAGGAGTGTGGCTTACCCATGGGTTGGTGTGAGGATTCAAGACTGGATGGATA
 AACTTAGGCAAGATCCGGCACACCATGGGGCTGGCTGGTCCCTGGCTGGTGAAGG
 ACTTGGCTGCCCTCCCACTCACACCCCTGGGTTCTGCCCTCCGGCTCTCGCAG
 GTTCCCCACCCGGTGTGCCCTGGTCACTGAGGACATGATAGGCACAGCCTT

12302 AGCCCCACATAACCTATGGGAGAGGTTACTAAACTTTCTAACGGTGTGAAACCAA
 GGCTCAGAATGGTTAAGTAAATTGTCACAGGGCCACAGAGGTAGGGTAGGTCTGG
 ATTAAAACTCCAAGTCTGGACTCCAGACCTCTAGGCTGTACTGTCATAGGGAGGA
 GTCTCACCCACCTAGGGCAGAGAAGAAAATCTTAAAGCCAGAGAAGTGTGGCTCATC
 TGTGGTCACCCAGAGAGACAGTGTGAGGACAGGGAGAAAATTATAACCTCAGTTCCAG
 [C, G]
 CCAAGGATCTGCTTGTACATAACCCAAACAGCCCCGCTATGGTGGTATTTCTTAGGT
 TCATATGGCGGCTTTGTTCCATTGATCTTCACAGCAATTCTCTACAGGAATCTGGC
 AGATTTATTCCTTAAAGGAATTCCAGGTCTTAAATCTATAGGGGCAACTATCAA
 ACTTACCCAAATGTTGGCCCTACCCACACACAAACAGGCCAGGGCGATCAGAAAG
 CACTGCTGAGCTCTGTCAGGGCCCACGCACTGCTGTGAGACAGAGAGGGAACTCA

12354 GAAACCAAGGCTCAGAATGGTTAAGTAAATTGTCACAGGCCACAGAGGTAGGGAGGT
 GAGTCTGGATTAAAACCTCAAGTCTGGACTCCAGACCTCTAGGCTGTACTGTCATAG
 GGAAGGCACTCTACCCACCTAGGGCAGAGAAGAAAATCTTAAAGCCAGAGAAGTGA
 GGTCTCATCTGTTGTCACCCAGAGAGACAGTGTGAGGACAGGGAGAAAATTATAACCTCA
 GTTCCCAGCCAAGGATCTGCTTGTACATAACCCAAACAGCCCCGCTATGGTGGTATT
 [C, T]
 CCTTAGGTTCATATGGGGCTTTGTTCCATTGATCTTCACAGCAATTCTACAGGA
 ATCTGGCAGATTATTCCTTAAAGGAATTCCAGGTCTTAAATCTATAGGGGCAA
 CTATCAAACCTCACCAATGTTGCCCTACCCACACACAAACAGGCCAGGGCGA
 TCAGAAAGCACTGCTGAGCTCTGTCAGGGCCCACGCACTGCTGTGAGACAGAGAG
 GGAACTCACATTATTGATCACCTACTGAGCATCATCACTAGGCTAGGACCGTCACATT

12487 ACCCACCTAGGGCAGAGAAGAAAATCTTAAAGCCAGAGAAGTGTGGCTCATCTGTGG
 TCACCCAGAGAGACAGTGTGAGGACAGGGAGAAAATTATAACCTCAGTTCCAGCCAA
 GGATCTGCTTGTACATAACCCAAACAGCCCCGCTATGGTGGTATTTCTTAGGTTCAT
 ATGGCGGCTTTGTTCCATTGATCTTCACAGCAATTCTCTACAGGAATCTGGCAGAT
 TTATTCTTAAAGGAATTCCAGGTCTTAAATCTATAGGGGCAACTATCAAACCTT
 [C, T]
 ACCCAATGTTGCCCTACCCACACACAAACAGGCCAGGGCGATCAGAAAGCACTG
 CTGAGCTCTGTCAGGGCCCACGCACTGCTGTGAGACAGAGAGAGGGAACTCACATT
 ATTGATCACCTACTGAGCATCCATCACTAGGCTAGGACCGTCACATTCTTAACTTTGA
 ATCCCTTACAGGGTAGGCATTATTCTCTTGTGACATAGCCATTAAAGAACAA
 AAATTGGGGCTGGGTGACTCACACCTGTGACTAGCACTTAAAGGGCTGAGGC

13198 CTAACATTCAAGGAAGGTTAGGCCGGAGCACAACATTGGGTTCCAGGGTTGAGGCTCCAG
 TGAGCTGATCTGCCACTGCACTACAGCCTGAGCAACAGAGCAAGACCCGTGACTCCAA
 AAACAAACAAACACACATTGAAACCAACAGATCTGACCCAAAGATGATGCTTAA
 TAGATGCCACCTCCCTGTGCTGGGCTTCACTAAAACACAGACAGATCAGGCAAC
 CAAGTCATCTAAGGAAAGAGGAAAGTGTAAACCAAAAGCACAAATACATAAAATTGCA

FIGURE 3M

		[- , A]
		AAAATGCTATTTAAAGAAAAAAAAGAGAAGAGAGGGCTTGAGGTTGACTAACAGAGAAAT GCCCTGGCTAATCCAGGAAGACTCCTGAAAGAGGTTGTTTTCCCCAGGTCTGCTT TGACATCTCTTTCACAGTGATCTGGTAGTGAGCTCCTCTCCCTCTCCAGA GCCTGCCCTATGGTGGCAGTGGGTGCGCTTCTCCGTCTGGTGGCTTCCAGA CTCAGTTGTAAGTGATAGCTCCGCCCTCTAGGCCACAGTCGGTCCCTGGGCCAGCC
13257		GTGAGCTGATCTGCCACTGCACTACAGCCTGAGCAACAGAGCAAGACCCCTGTGACTCCA AAAACAAAACAAAACACATTGAAACCCAAACAGATCTGACCCAAGATGATGATGCTT ATAGATGCCACCTCCCTGTGCTGGGCTTACTAAAAACACAGACAAGATCAGGAA CCACAGTCATAAGGGAAAGAGGAAAGTGTAACCAAAGCACAAATACATAAAATATTG CAAAAATGCTATTTAAAGAAAAAAAAGAGAAGAGAGGGCTTGAGGTTGACTAACAGAGA [A , G] TGGCCTGGCTAATCCAGGAAGACTCCTGAAAGAGGTTGTTTTCCCCAGGTCTGCTT TTGACATCTCTCTTTCACAGTGATCTGGTAGTGAGCTCCTCTCTCCTCTCCAG AGCCTGCCCTATGGTGGCAGTGGGTGCGCTTCTCCGTCTGGTGGCTTCCAG ACTCAGTTGTAAGTGATAGCTCCGCCCTCTAGGCCACAGTCGGTCCCTGGGCCAGC CGCAAAGGGCTTCATGCCACGGCTGGCTAGTCCACTGTACCTCCACCTCTGGGCC
14541		TCATGGACACTGACATTATGTGAATCCAAGACCTATAATAGGGTAGGTAAATTCAAGCT TATGACCTCCTTCTTGTCTGACCCACCCAAAGAAGAGGTTGCTTTAAAGCCAATA AAGACATTCTGCAACTTGAGCTCAGTCCCTGTACAGGCCAGGATATCCAGGGAT TAAAATCATCACGTACTGCTCCCTCTACTTTGCAACTCAGAGATCTTCAGGCAAA GGTCAATGCCAAGGTAAAGGCTCAGTCCCTGGCGACCAGAGGCTCTGGACAGAGTGGCC [G , A] GAAAATGGAAGCAGAAGGGCGGTGGGAGCTGAGAATAGGCCACTCCATAGAGGGGGAG GTCAAGATTGCTTGGCTCTCCCTGCAAGACAGGCATGGACCCCAGAAAGTATTACT AGCCAAGAAAAATACCTCAAGAACAGGAGAAGCGGAGATGAGGCCACACAACAGAG GAGGTCTTATTCTGAAACCAAGGTGAATGAAGGCCAGAAGCAGCCCCGTGCCCTGCT CTCCGCCATTCTGATACTGCCCCCTGTTACTCATGGTACCCCTGGGGGCCCCGCTTCCC 14545
		GGACACTGACATTATGTGAATCCAAGACCTATAATAGGGTAGGTAAATTCAAGCTTATG ACCTCCTTCTTGTCTGACCCACCCAAAGAAGAGGTTGCTTTAAAGCCAATAAAGA CATTCTGCAACTTGAGCTCAGTCCCTGTACAGGCCAGGATATCCAGGGGATTTAA ATCATCACGTACTGCTCCCTCTACTTTGCAACTCAGAGATCTTCAGGCAAAGGTC ATGCCAAGGTAAAGGCTCAGTCCCTGGCGACCAGAGGCTCTGGACAGAGTGGCCGGAA [A , G] ATGGAAGCAGAAGGGCGGTGGGAGCTGAGAATAGGCCACTCCATAGAGGGGGAGGTCA AGATTGCTTGGCTCTCCCTGCAAGACAGGCATGGACCCCAGAAAGTATTACTAGCC AAGCAAAATACCTCAAGAACAGGAGAAGCGGAGATGAGGCCACACAACAGAGGAGG TCTCTATTCTGAAACCAAGGTGAATGAAGGCCAGAAGCAGCCCCGTGCCCTGCTCTCC TGCCCATCTGATACTGCCCCCTGTTACTCATGGTACCCCTGGGGGCCCCGCTTCCCACCC
15041		ACCAAGGTGAATGAAGGCCAGAAGCAGCCCCGTGCCCTGCTCTCTGCCATTCTGATAC TGCCCCCTGTTACTCATGGTACCCCTGGGGCCCGCTTCCACCCCTGACAGGCAAAGACA GAAAGTCTGGAAACACTGCCGGTGGCCGCTGGCATTTTCTTTCTTTCTTTCTTT TCTTTAGAGATGGAATTGTCTTGTCACCCAGGCTTGAGTGCAATGGCGTTATCTT GGCTCACTGCAACCTCCACCTCTGGGTTCAAGCGATTCTCTGCCCTAGCCTCCAAAGT [C , A] GCTGAGATTACAGGTGCCACCAACCCCAGCTAATTTTGTTAGTAGATATTGGGT TTCAACATGTTGGCCAGGCTGGTGCAAAACTCTGACCTCAGGTGATCCACCTACCTTAG CCTTCAAAGTGCTGGGATTACAAGCCTGAGCCACTGCCAGCCTGGCATTTTCTT CTTGGATGAGGTGCTACCATCTCCAGGAAAGCCACTGAACCCCCAAGGCCCTTCCAT TTTCTGGCTAAGATAGGACATGGACTTTGAACAACCCAGAGGGGAAACAGCA
15053		GAAGGCCAGAAGCAGCCCCGTGCCCTGCTCTCTGCCATTCTGATACTGCCCTGTTA CTGATGGTACCCCTGGGGCCCGCTTCCACCCCTGACAGGCAAAGACAGAAAGTCTGG GAACACTGCCGGTGGCCGCTGGCATTTTCTTTCTTTCTTTCTTTCTTTAGAGA TGGAAATTGTCTTGTCACCCAGGCTTGAGTGCAATGGCGTTATCTGGCTACTGCAA CCTCCACCTCTGGGTTCAAGCGATTCTCTGCCCTAGCCTCCAAAGTCGCTGAGATTAC [A , C] GGTGCCACACACCCAGCTAATTTTGTTAGTAGATATTGGGTTTACCATGTTG GCCAGGCTGGTGTCAAACCTCTGACCTCAGGTGATCCACCTACCTAGCCTCCAAAGTG CTGGGATTACAAGCCTGAGCCACTGCCAGCCTGGCATTCTCTGGATGAGGT GCTACCATCTCCAGGAAAGCCACTGAACCCCCAAGGCCCTTCCATTCTGGCTAAG ATAGGACATGGCCATGGACTTTGAACAACCCAGAGGGGAAACAGCAGTGAATTCTG
15065		CAGCCCCGTGCCCTGCTCTGCCATTCTGATACTGCCCTGTTACTCATGGTACCC

FIGURE 3N

TGGGGGCCCGCTTCCCACCCGTACAGGCAAAGACAGAAAGTCTCTGGGAACACTGCCTG
 GTGGCGCTGGCATTTCCTTTCTTTCTTTCTTTAGAGATGGAATTGCT
 CTTGTCACCCAGGCTTAGTGAATGGCCTTATCTGGCTCACTGCAACCTCCACCTCTG
 GGTTCAAGCGATTCTCTGCCCTAGCCTCCAAAGTCGCTGAGATTACAGGTGCCACAC
 [A, G]
 CCCAGCTAATTTCGTATTTTAGATATTGGTTTACCATGTTGGCCAGGCTGGTG
 TCAAACCTCTGACCTCAGGTGATCCACCTACCTAGCCTCCAAAGTGTGGGATTACAA
 GCCTGAGCCACTGCGCCCAGCCTGGGCTTTCTCCATTTCCTGGATGAGGTGCTACCATCTCC
 CAGGGAAAGCCACTGAACCCCCAAGGCCCTCTCCATTTCCTGGTAAGATAGGACATGGC
 CCATGGACTTTGAACAACCCAGAGGGGAAACAGCAGTGAATTCTGGGAACCCAGGC

15108 TGTTACTCATGGTACCTGGGGGCCCGCTTCCCACCCGTACAGGCAAAGACAGAAAGTC
 TCTGGGAACACTGCCTGGTGGCCGCTGGCATTTCCTTTCTTTCTTTCTTTCTTT
 AGAGATGGAATTTCGTCTTGTACCCAGGCTTAGTGAATGGCCTTATCTGGCTCAC
 TGCAACCTCCACCTCTGGGTTCAAGCGATTCTCTGCCCTAGCCTCCAAAGTCGCTGAG
 ATTACAGGTGCCACACACCCAGCTAATTTCGTATTTTAGATATTGGTTTCAAC
 [A, C]
 TGTTGGCCAGGCTGGTGTCAAACCTCTGACCTCAGGTGATCCACCTACCTTAGCCTTCA
 AAGTGCTGGATTACAAGCTGAGCCACTGCGCCCAGCCTGGCATTTCCTGGAT
 GAGGTGCTACCATCTCCAGGGAAGCCACTGAACCCCCAAGGCCCTCTCCATTTCCTGG
 CTAAGATAGGACATGGCCATGGACTTTGAACAACCCAGAGGGGAAACAGCAGTGAATT
 TCCTGGGAACCCAGGCCAGGGCTAGCAAGGCTGGGTGCCATGGCAGTAATCCT

16274 CTTCCAGACTGTCTCCCTGCAGGAGCTGCAGCAGGACTTGGAGATGCGCCCCCACC
 CCCCCAACACAACACCAGACCCCGCTAACGGCACCGCAGCGTGTCTTATACACCTTCAG
 TGACAGCTCTCACCTGCCAGAGTGAGCCACCAGCCTCGTGGGCCCCGGCGAGCC
 CAGTGACATGCTGGCCAGCGTCCCACCCCTCGTACCTTCCACACCCATCCTGGACAT
 GAGTGGAGTCAGCTCGTGGACTTGTGGCATCAAGGCCCTGGCCAAGGTGAGGCCCTC
 [-, G]
 GGGACAGCAAGCACCACCCACTCCACCCCCCTCGCTCTCCACATTCCCTTCTG
 GGAGCCCTCATTTAGGAAGCTGAGGGAGGAAGCTACTGGGAGACTAACAGCTCTAG
 GAATCCCTCTTCCCCAGGCCACAGGTTGAGACATTCTCCACAGAGCAGGGCCAGA
 CGGCCCATGACAATGAGTGGCGGACAAGTCTACCAGAGTTCAAGGCCCTGTGCTCCCA
 ACACCCCCAGCAGTGGCATCCAAAGTCCCTCAGCCATCAGGAACCCACCCAGGTTCT

17424 AACATGGTAAACCCCCCCTACTAAAAATACAAAATTAGCCAGGTGTGGTGACGGGC
 CCCTGTAGTCCCAGCTACTCGTAGGCTGAGGCAAGAATTGCTTGAACCCAGGAGGCGG
 AGGTTGAGTGGCAAGATCGGCCACTGCACTCCAGCCTGGCAACAGAGTGAACACT
 CATCTAAAGAAAAAAAGAAAAATCTAGCCCCACAGAAGGGCCATGGTACTTT
 AAGTGCCCGCACGTTGCAAAGTCCATTCCGCTCCACTTCCAGAGAAACCGTCAGC
 [C, T]
 AACACTCCAGGGAGAAGTGGTGTGCTTGTCTATTTCGTCTGGCTGCTGGCT
 CAGGGTTGCTTATTTCGTCTGGCTTCCCTCTGAAGTACGTTGTGAATCACTTTGAGA
 CCCACTGACAACATTCTTCCCTTGGCTCCCTACCCAAACACCTCTAGCTGAGCT
 CCACCTATGGGAAGATGGCGTGAAGGTCTTGTGGTGAACATCCATGGTAAGAGAAAGA
 GGACATTAGGGACTGAAAGACTGGCAAGGAGTGTGGGTAGGAACAGGTTGGGGTC

17627 AATATCTAGCCCCACAAGAAGGGCCATGGTACTTAAAGTGCCGCCACGTTGGCAAA
 GTCCATTCCGCTCCACTTCCAGAGAAACCGTCAGCCAACACTCCAGGGAGAAGTGGT
 TGCTTGTGCTATTTGTCTTGGCTGCTGGCTCTCAGGGTTGCTTATTGTTGGC
 TTCCCTCTGAAGTACGTTGTGAATCACTTGGACCCACTCAGAACATTCCCTTCC
 TTTGCTCCCTACCCAAACACCTCTAGCTGAGCTCCACCTATGGGAAGATCGCGT
 [G, A]
 AAGGTCTTCTGGTGAACATCCATGGTAAGAGAAGAGGACATTAGGGACTGAAAGACT
 GGCAAGGGAGTGTGGGTAGGAACAGGTTGGTGGGCTGTAATAGTGAGGAGGTTGGAAAC
 GAGAGCACCCAGTATCCCCACAAGCTGCTGCTGCTCATAAAGCTTCAGGTACAAGT
 CCAAAGAGACTGGTCAAGATTGATAAACATCCTAGGGCCTTAGTGAAGAGTGGGGTG
 AGGAGGTATGGAGTTACAGAAGGACAGCTAGGATTCTAATCTACCCATAACTAATTG

18427 GGGTGCATATACACAGCCTCAAGGACGTGGCACAGGGCAGCAGACATTACATGACTAG
 CATGTACGAAAGTGCAGAGATGTGGAGCAAGTGCACACAGACACAGGAGAATGTGA
 AGGGGCACATACACACACACACCCAGCTCCCTGCACTGGCTCAGACCCCCCTCAGCAGGGCT
 GCAGTTCCAAGCTCCGATGGCACGTTGGGAGAGAATCTGCACTGGCAATGACCTG
 CTATGATATGTTCTGGAGTTAGAAGCAGTGGATTCTCCAAACCTCACTGGACACCCCT
 [T, C]
 AGGAAACCATCTCTAGGATTAAGAGTAATCCACACAAACTTCAATGCCACACATTGGAA
 GTTGCTGGAAAGGCTGGAAAACAAGAGGAAGGATGGGTCTTGGGGATAGAAGTGGC
 AGCGGCCTTCAAGGATGGCTTAGGCTTCCACTCGAATCACCACAAAGTACTGACTC

FIGURE 30

FIGURE 3P

		[T, G] AACAGATTAGAAAATATTCAGGAGGACAGGAATTCCCAAGGTCAAGGCAGCTAGCC AATAGTTTCTAAGCTGAGTAAACCTTCCCTGCCTAAACGGCCACAAAGGAGGGAA GACCGCGATACACACCTGCTGGTATAAGGGGAAGACACAGCCGTGCTGTTTGAG GGCAGGTAAAGGAAGGGCAAGAGGATAAGTCATGTGTCAGGAAGCAGCGTCAACCAGA GCCGGCCACCTGTCCTTCTGCCACCATGCACCAACTTGCTGTTCACTGAAG
20180		TTCTGCTTCGTGAAATTGTCAGGCAAGAGGGAGATTCTGGAGGAAGCTGCTGATT AGTTGTTAGTGCCTAATCATGTCAGTACTCTAGTTGTTGATGTTACTTGATTAGTC TAGCACTTATAAAATAATTATATTATATAATATACATCATATTAGACCAATTCA CAGATACAAATCACACACATAAAACACACACCTTCAACAGCATTGAGGGACAAAGCA GGCAAGTGGGCTGGTTACAGACTTAAACAGATTAGAAAATATTCAGGAGGACA [G, A] GAATTCCCCAAGGTCAAGGCAGCTAGCCAATAGTTTCTAAGCTGAGTAAAACCTTCCCT GCCTCTAACGGCCACAAAGGAGGGAAAGACCCGCGATACACACCTGCTGGTATAAGGGGG AAAGACACAGCGTGTGTTTGAGGCGAGGTAAAGGGAAAGGGCAAGAGGATAAGTC TGTGTCAGGAAGCAGCGTCCAACCCAGCGGCCACCTGTCCTTCTGCCACCATGC ACCAACTTGTGTTCACTGAAGCTTCTGCACTGGCTTCTCCCTCCAGGCT
20584		TGTCTGGTATAAGGGGAAGACACAGCCGTGCTGTTTGTGAGGCAGGTAAAGGGAGG GGCAAGAGGATAAGTCATGTGTCAGGAAGCAGCGTCCAACCAGAGCCGGCACCTGTC TTTCTGCCACCATGCACCAACTTGTGTTGCTGACTGAAGCTTCTGCACTGGC TTCTCCCTCCAGGCTCAGGGGATGCTGAGCTCTCTGTACGACTCAGAGGAGGACA TTGCGAGCTACTGGGACTTAGAGCAGGTGAGCTGAGGGAAAGGGCTGTGAGGGTGGAGC [A, T] GGGCGAAGAGGGGAAGGGATGGGTCGCTGTCAAATACAAGGCCTTCACTCAGCTGTC CCTCCAGCCCAGAGCAGTCACATTCAAGGCCACAAAGATTGTGGTCACTTTGTTT TTCTTTCTTTCTTTTTTTTTAATTGAGACAAAGTCTCACTCTATCACCC AGACTGGAATGCACTGGCATGATCTCAGCTCACTGCAACCTCTGCCCTCCGGTTCCAGA GGTTCTCTGCCCTCAGCTCCCAGTAGCTGGACTTCAGGCGTGCAGCCAGCTAATT T
20717		ATGCACCAACTTGCTGTTCACTGAAGCTTCTGCACTGGCTTCTCCCTTCCA GGCTCCAGGGATGCTGAGCTCTCTTGACACTCAGAGGAGGACATTGCACTGACTG GGACTTAGAGCAGGTGAGCTGAGGGAAAGGGCTGTGAGGGTGGAGCAGGGCAAGAGGG GAAGGATGGGTCGCTGTCAAATACAAGGCCTTCACTCAGCTGTCACCTCCAGGCCAG AGCAGTCACATTCAAGGCCACAAAGATTGTGGTCACTTTGTTTCTTTCTTT [T, C] CTTTTTTTTTTTTTAATTGAGACAAAGTCTCACTCTATCACCCAGACTGGAATGCA GTGGCATGATCTCAGCTCACTGCAACCTCTGCCCTCCGGTTCCAGAGGTTCTCTGCC CAGCCTCCCGAGTAGCTGGACTTCAGGCCCTGCGCCCAGTAATTGTTGTTAGTA GAGACAGCTTTTACCATGTTGGCTGGCTGCTGCAACTCCGATCTCAAGCAATCTG CCTGCCCTGGTCTCTAAAGTGCCTGGATTACAGGCATAAGGCCACGATGCCCTGGCTTGT
20894		GGGGAGGATGGGTCGCTGTCAAATACAAGGCCTTCACTCAGCTGTCACCTCCAGCC CAGAGCAGTCACATTCAAGGCCACAAAGATTGTGGTCACTTTGTTTCTTT TTTCTTTTTTTTTAATTGAGACAAAGTCTCACTCTATCACCCAGACTGGAA TGCACTGGCATGATCTCAGCTCACTGCAACCTCTGCCCTCCGGTTCCAGAGGTTCTC GCCTCAGCTCCCGAGTAGCTGGACTTCAGGCCCTGCGCCCAGTAATTGTTGTT [A, G] GTAGAGACAGCTTTTACCATGTTGGCTGGCTGGTCTGCAACTTCCGATCTCAAGCAAT CTGCCCTGCCCTGGCTCTAAAGTGCCTGGATTACAGGCATAAGGCCACGATGCCCTGG TGTGTTCTTCTCACTCCCTGAAAGGCATCGTGGGAGAGGGTGAAGTCAGTGGACCA AGTCCTAGAGAACAGTATCTATTCTCAACACATCACCCACGTGACCCCTGAGC AAGGCCACATACACCTGGGCCAGTAGTTTATCATCTGAAATTAGGGAAACATAGGT
21787		GGGTGCAATGGTTCACACCTGTAATCCAGCACTTGGGAGGCTGAGGTGGCGGACAC CTGAGGTCAAGGAGTTGAGACAGCCTGCCAACATGGGAAACCCCGTCTACTAAAA GCACAAAATTAGCCAGCGTAGTGGTCACTGCCCTGAGTCCCAGCTACTCGGAAGCTG AGGCATGAGAATCACTGAACTGGAGGAGATGTTGCACTGAGGCCAGATCGTGC TGCACTCCAGCTGGGTACAGAGCTAGACTGTCACAAACAAACAAACAAACAAAC [-, A, C] TAAAGATATGGAATATGAGGGATCACCACCCATAGGGCCCTGGATTAAACACCA CCACCAATGCCCTGAATTAAAGAAACAGATGACTAGGTTGGAGAAATCTGGTTGG GTCTATGAGAAGTAGTGTCTCTTGTGCCCTTCCCAATTCTTTGACATTGAGCTC ATGGTGCTGTCATCCGTCACAGTGTGATGGCAGGTGGAGACAGATTAGAAAATAG AGCTGGAGGCCACAGAGATTGGCAGACTGATTGCGCTCTTGGAAATCTCCAGCACA
22264		CTCCATGGTGCTGTCATCCGTCACAGTGTGATGGCAGGTGGACAGATTAGAAA

FIGURE 3Q

ATAGAGCTGGAGCCACAGAGATTTGGCAGACTGATTCGGTGCCTCTTGAATCTCCAG
 CACATTCCAAAAGCCTGGATAGGACAAAATAGCTTATCAACGTGAGAAAGGACTTCAG
 AGCTTGCTACTGCCAACCTCATTTACCAATGAGGAAAGTGAAGCTATTAGGGGGCG
 AGGGACACGTGGAAGGTACACAGCACACAGGAGGTGATTACATGTAGATTCAGCACC
 [T, C]
 GCTCCTGCCACGCTGGACTGGTTACCTCTAGGCTGACCCCTGCCCTCCCCCTGTCACA
 CACACTCTGCACACACACACACACACACACACAGGTGCTTGTCTGCCAGG
 GGTTCTAGGGTACCTCTGGTTGCAGCCACTGTGACCCCAACTGGCTAACCTCTC
 TCCCCCTCCACTTCCCTGTGGTCTGCAGGAGATTCAGGGAGCATGTTACAGC
 AGAGACCCCTGACCGCCCTGTGAGGGCTCAGCCAGTCCTCATGTCCTACAGAGTGCCTG
 22338 ACAGAGATTTGGCAGACTGATTCGGTGCCTCTTGAATCTCCAGCACATTCCAAAAG
 CCTGGATAGGACAAAATAGCTTATCAACGTGAGAAAGGACTTCAGAGCTTGCTACTGC
 CAACCCCTCATTTACCAATGAGGAAAGTGAAGCTATTAGGGGGCGAGGGACACGTGGAA
 GGTACACACAGCACACAGGAGGTGATTACATGTAGATTCAGCACCTGCTGCCACGC
 TGGACTGGTTACCTCTAGGCTGACCCCTGCCCTCCCCGTTCACACACACTCTGCAC
 [-, C, A]
 CACACACACACACACACACACACAGGTGCTTGTCTGGCCAGGGGTTCTAGGGTCA
 CCTCTGGTTGCAGCCACTGTGACCCCAACTGGTCTAACCTCTCTCCCCCTCCACTT
 CCTCCTGTGGTCTGCAGGAGATTCAGGGAGCATGTTACAGCAGAGACCCGTGACCG
 CCCTGTAGGGCTCAGCCAGTCCTCATGTCCTACAGAGTGCCTGGACTGGACTTC
 CATAAAGGATGAGCCTGGGTACAGGGGGTGTGGCGAGGAAAGTGCATCCCCCAGA
 23363 CAGGGACCATGTGCTCCACACCCAGGAGTCTAGGCCTGGTAACTATGCGCCCCCGT
 CCATCATCCCCAAGGCTGCCAACACCACACTGCTGTAGCAAGCACATCAGACTCTAGC
 CTGGACAGTGGCCAGGACCGTCGAGACCAAGAGCTACCTCCCCGGGACAGCCACTA
 AGGTTCTGCCCTAGCCCTGAACACATCACTGCCCTCAGAGGCTGCTCCCTCCCCGTGA
 GGCTGGCTAGAAACCCAAAGAGGGGATGGTAGCTGGCAGAATCATCTGGCATCTAG
 [T, C]
 AATAGATACCAGTTATCTGCACAAAACCTTGGAACTTCTCTTGACCCAGAGACTC
 AGAGGGGAAGAGGGTGTAGTACCAACACAGGGAAACGGATGGGACCTGGGCCAGACA
 GTCCCCCTTGACCCCAAGGGCCATCAGGGAAATGCCCTTGGTAATCTGCCCTATC
 CTTCTTACCTGGCAAAGAGCCAATCATGTTAACCTTCTTATCAGCCTGTGGCCAGA
 GACACAATGGGTCTTCTGTAGGCAAAGGTGGAAGTCCTCCAGGGATCCGCTACATCCC
 23688 AAACTTTGGGAAATCCTCTTGCACCCAGAGACTCAGAGGGGAAGAGGGTGTAGTACC
 AACACAGGGAAAACGGATGGGACCTGGGCCAGACAGTCCCCCTGACCCAGGGCCAT
 CAGGGAAATGCCCTCTTGGTAATCTGCCCTATCCTTACCTGGCAAAGAGCCA
 TCATGTTAACCTTCTTATCAGCCTGTGGCCAGAGACACAATGGGGTCTTCTGTAGG
 CAAAGGTGGAAGTCCTCCAGGGATCCGCTACATCCCCTAACATGCACTGCAAGATGTGGAAAG
 [G, A]
 GGCTGATCCAGATTGGGCTTCTGCACAGGAAGACTCTTAAACACCTTAGGACCTCAG
 GCCATCTTCTCTATGAAGATGAAAATAGGGGTTAAGTTCATATGTACAAGGAGGTA
 TTGAGAGGAACCCCTACTGTTGACTTGAAGGAAATAAATAGGTTCCATGTGAAAGTGTGTTGTA
 AAATTCAGTGGAAATGACAGAAAATCTCTGGCTCTCATCACTGCTTTCTCAAGCT
 TCTTCAGCTTAACAACCCCTTCCCTAACAGGTTGGCTGGCCAGCCTAGGAAAACATCC
 24210 TCACTGCTTTCTCAAGCTTCAAGCTTAACAACCCCTTCCCTAACAGGTTGGCTGGC
 CCAGCCTAGGAAAACATCCCATTCTAACCTCAGCCAGACCTGCCTGTGTCTGTGT
 GTTGAGTGTAGCTGGTCACTAACAGTCTTCTAGAGTTAAAGGAGGGGGTGTGGCAA
 GAGCCAACACATTCTGGCCAGGAGCATTGCTTCTGTGAATTCAATTGCACTCTGG
 CTGCCAATGGAACTCAAAACTTGGAAAGGCGAAGGACAATGTTATCTGGATTACCGTGC
 [A, C]
 CAGCACCCGAAGTGCCTAAATTCCAGGGAGCAAGAGGCCTAGCCAATGACAACACTCACTCT
 CCCCTACTCCACCTCTTCAAGTCAGGCTCAGGCCCAGGAGGTGGGAGAAGGTCAAGA
 GCCTCAGGAATTCTCAAGTCAGAGTCCCCCTTGAACCAAGTATCTAGATCCCCTGAGGAC
 TTGATGAAGTGTCTTAACCCCCAAGTAATCTTAACCCCCAGACCCCTCAGAACTG
 AAGGAGATTGTTGACCCAGTGACCTGGAGTTGAGGCTCAGGGAGAGATCTGCCACATGTC